Structural and functional studies of the budding yeast kinetochore complex CBF3.

Vera Leber

University College London

and

The Francis Crick Institute

PhD Supervisor: Martin Singleton

A thesis submitted for the degree of

Doctor of Philosophy

University College London

December 2018

Declaration

I, Vera Leber, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Kinetochores connect centromeres with spindle microtubules during mitosis and meiosis to ensure correct chromosome segregation. The initial step of kinetochore establishment in point centromere species, like budding yeast, is the sequence-specific recognition of the centromere by the essential CBF3 complex. CBF3 consists of four proteins, Ndc10, Cep3, Ctf13 and Skp1, and its binding to the centromere is required for the assembly of all other kinetochore proteins. Beside this, the complex is involved in the recruitment of the centromere-specific histone Cse4 through its interaction with the histone chaperone Scm3 and it also impacts the overall conformation of the centromere through DNA looping and bending. Besides its crucial role in kinetochore establishment, little structural data is available for the CBF3 complex, mainly due to difficulties with recombinantly expressing and purifying the full complex. This thesis describes a working co-expression and purification protocol for CBF3, which allowed structural as well as functional studies of the complex and led to the cryoEM structure of the core CBF3 complex, comprising the centromere-binding Cep3 and the regulatory subunits Ctf13 and Skp1. Besides the overall core architecture, this structure provides the first insights into the inherently unstable subunit Ctf13, as well as a potential new conformation of the Skp1/F-box interaction. Furthermore, it provides interesting insights into the unusual DNA-binding properties of the dimeric Cep3 to a single consensus site in the centromere, a matter of some debate in the field. Biochemical studies revealed a potential and previously undescribed regulatory mechanism of DNA-binding activity through phosphorylation of the Skp1 subunit. Additional work was undertaken to better understand how Ndc10 binds to the core complex and therefore how the full CBF3 assembles, as well as interaction studies with the centromere-specific nucleosome. As a side project, interaction studies between Ndc10 and Scm3 were undertaken, to elucidate the proposed Cse4-loading function of CBF3.

Impact Statement

Faithful chromosome segregation is a critical step in a cell's life and mistakes in this complicated process can lead to cell death or chromosomal instability, the latter being a hallmark of cancer. Understanding of this crucial process is therefore important, and whilst intensive previous research has shed light on the regulatory mechanism, proteins involved and their functions, lots still remains unknown. Budding yeast has served as a faithful and widely used model organism for kinetochore studies, as it comprises a much simpler but highly conserved environment. Although CBF3, the subject of this thesis, has no homologue in higher eukaryotes, its function may well be conserved. Already it is known, that it impacts the loading of the yeast homologue of CENP-A through its interaction with the conserved centromere-specific histone chaperone Scm3. However, its centromere-binding function might also be conserved, given its ability to change the overall conformation of the centromere by bending, which has also been proposed for the human centromere protein CENP-B. The role of CENP-B, despite being among the first centromere proteins to be discovered, is still unclear. Besides these functional similarities, CBF3 plays an essential role in budding yeast kinetochore establishment and therefore is crucial to ongoing attempts to reconstitute the whole budding yeast kinetochore. The latter will provide an invaluable tool to study kinetochore function and regulation and cannot easily be done with the more complex human kinetochore. CBF3 has so far been difficult to recombinantly express and purify, and was therefore lacking in the reconstituted kinetochores. The protocol described in this thesis should prove to be beneficial for future efforts in reconstitution of the whole budding yeast kinetochore, as well as for structural studies of larger parts of the inner kinetochore and centromere.

Publications arising from this work

Leber, V., Nans, A., and Singleton, M. R. “Structural basis for the assembly of the CBF3 kinetochore complex.” The EMBO Journal 2018, 37: 269-281

Acknowledgement

I would like to express my sincere appreciation to the many people contributing to my PhD study:
First of all, to my supervisor Martin Singleton for being a great mentor, a patient teacher and generally a respectful person and group leader, all of which deserves the highest appreciation in the present high-pressure scientific environment.

Besides my supervisor I would like to thank my thesis committee, Alessandro Costa, Caroline Hill and Kazunori Tomita for insightful meetings and comments. Especially, I would like to thank Alessandro Costa for all his help in cryoEM methodology.

A special thank you also goes to my fellow lab members for teaching me a lot, as well as for insightful discussions, but mostly for making all the years of work a great time and lots of fun!

Furthermore, I would like to thank all the STPs for their great services, especially Raffa Carzaniga for her invaluable help with the electron microscope and all other great advice and insight. Also, to David, Aaron and Colin from the Proteomics STP for the many mass spectrometry projects they carried out for me, to Andy for sharing his expertise in crystallography and to Vicky, Ramin and the rest of the sequencing facility for setting up countless sequencing reactions.

Equally, I would also like to thank the many laboratory assistants and managers for their help and support, especially Kim, Diaz, Paul, Karen and Donna. Without them I could not have managed all the work done for this thesis in the given time frame.

I also thank Peter Rosenthal, Phil Walker and Andrea Nans for Krios access, data collection and help in processing; Tillmann Pape for his help in screening cryo grids; Celine Bouchoux and Deniz for their help with yeast expression and *in vivo* experiments, and Paolo, Ferdos, Oli and Donald for their practical help and insight into cryoEM.

Last but definitely not least, a special thank you to my partner Matt for his support during good and not so good times and for patiently helping me to improve my English skills, as well as to my family and friends back at home for supporting me and my wish to do a PhD.

Table of Contents

Table of figures

List of tables

Abbreviations

Å	Angstrom (1Å=0.1nm)
APS	Ammonium persulfate
AUC	Analytical ultra-centrifugation
CATD	CENP-A targeting domain
CCAN	Constitutive centromere-associated network
CCD	Charge-coupled device
CDEI-III	Centromere DNA element 1-3
CEN3	Centromere of chromosome 3, *S.cerevisiae*
CENP	Centromere protein
CHIP	Chromatin immunoprecipitation
CIP	Calf intestine phosphatase
CoIP	Co-immunoprecipitation
COMA	Ctf19-Okp1-Mcm21-Ame1 (some of the subunits); or Ctf19 complex
CPC	Chromosome passenger complex
cryoEM	Cryo electron microscopy
CTD	C-terminal domain
CTF	Contrast transfer function
CV	Column volume
Da	Dalton
dsDNA	Double strand desoxyribonucleic acid
DSSO	Disuccinimidyl sulfoxide
dNTPs	Deoxyribonucleotide triphosphate
DQE	Detective-quantum-efficiency
DTT	Dithioreitol
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
EMSA	Electrophoretic mobility shift assay
6FAM	6-Carboxyfluorescein
FACS	Fluorescence-activated cell sorting
FBXLs	F-box proteins containing LRRs
FBS	Fetal bovine serum

FSC	Fourier shell correlation
Fw	Forward
GF	Gel filtration
GraFix	Gradient fixation
HFD	Histone fold domain
Hr	Hour
HRP	Horseradish Peroxidase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KMN	KNL1-MIS12-NDC80 network (*S.cerevisiae*: Spc105-Mtw1-Ndc80)
LB	Lysogeny broth
LiAc	Lithium acetate
LN_2	Liquid nitrogen
LRRs	Leucine-rich-repeats
Min(s)	Minute(s)
MIND	Mtw1 including Nnf1-Nsl1-Dsn1 complex
MSC	Multiple cloning site
MTF	Modulation transfer function
NS	No-salt
NTD	N-terminal domain
OD_{600}	Optical density at 600nm
ON	Overnight
ONC	Overnight culture
PEG	Polyethylen glycol
PBS	Phosphate Buffered Saline
PCR	Polymerase chain reaction
PTM	Posttranslational modification
RF	Restriction-free
RT	Room temperature
Rv	Revers
SAC	Spindle assembly checkpoint
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate - Polyacrylamide gel electrophoresis
Sec	Second

SEC-MALS	Size exclusion chromatography – multiple angel light scattering
SOC	Super optimal broth with catabolic repression
STP	Science technology platform
SCF ligase	Skp1/Cullin/F-box ligase
SAC	Spindle-assembly checkpoint
TE (buffer)	Tris-EDTA (buffer)
TEMED	Tetramethylethylenediamine
TBE	Tris-boric acid-EDTA
TEM	Transmission electron microscopy
YP-media	Yeast extract peptone media

Chapter 1. Introduction

1.1 Mitosis and chromosome segregation

One of the characteristics of a living organism is its ability to propagate. To be able to do so cells need to divide, which is a complicated process and needs to be heavily regulated to avoid abnormal growth, as happens in cancer (Hanahan and Weinberg, 2000). Before a cell divides, it needs to replicate its entire genome, which happens during S-phase of the cell cycle. The two identical copies must then be organised and separated into the two emerging daughter cells. This process, also called mitosis, was first described in the 18th century by a number of scientists, one of them being Walther Flemming, who coined the term. Mitosis is only a short period of the cell cycle and is preceded by interphase, in which the cell grows in size, replicates its genome and generally prepares for mitosis (Figure 1).

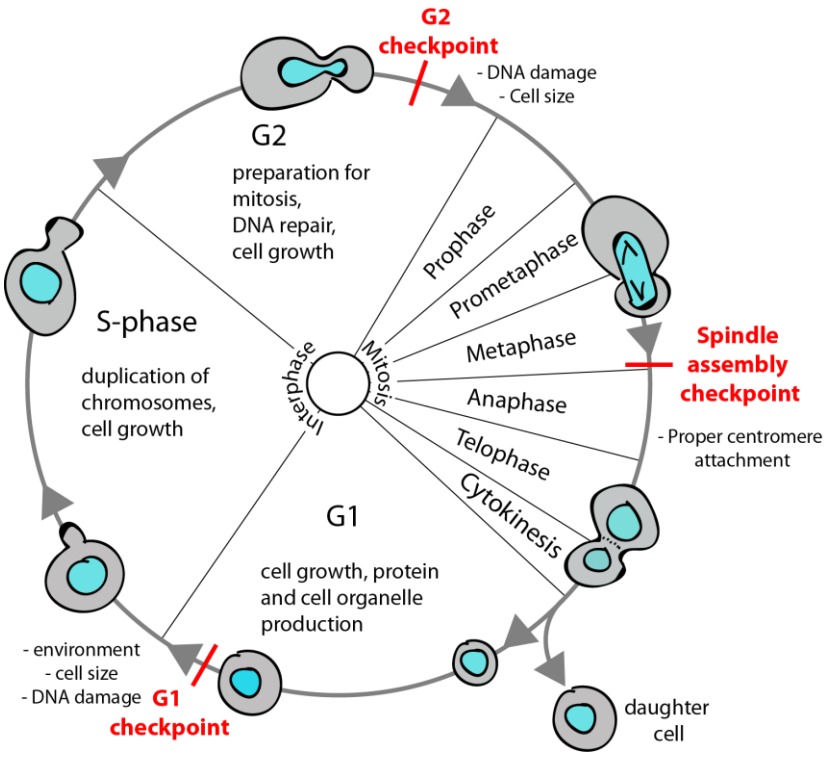

Figure 1: The cell cycle of budding yeast

Mitosis and cytokinesis is followed by the much longer interphase, which can be divided into G1, S-phase and G2. The progression through interphase is heavily regulated by cyclins and cyclin-dependent kinases and cells must fulfil checkpoints (highlighted in red) to progress into the next stages. In G1 the cell grows in size by normal metabolism. Replication of the genome happens in S-phase, followed by G2 in which the cell has time for DNA repair of damage happened during replication. If the G2 checkpoint is fulfilled it enters mitosis in which the sister chromatids are separated. Cytokinesis describes the process of the physical separation of the two emerging daughter cells and therefore the end of one cell cycle and beginning of the following one.

The process in which the mitotic spindle physically separates the two sister chromatids is called chromosome segregation (Figure 2). The enormous macromolecular complex, which facilitates this separation, is the kinetochore (McAinsh et al., 2003, Santaguida and Musacchio, 2009, Perpelescu and Fukagawa, 2011, Biggins, 2013, Westhorpe and Straight, 2013, Cheerambathur and Desai, 2014). The function of the kinetochore is to connect the centromeres of the sister chromatids to spindle microtubules of the opposite pole. Once this connection is correct, sensed by the kinetochore, the sister chromatids can be pulled apart into the two emerging daughter cells.

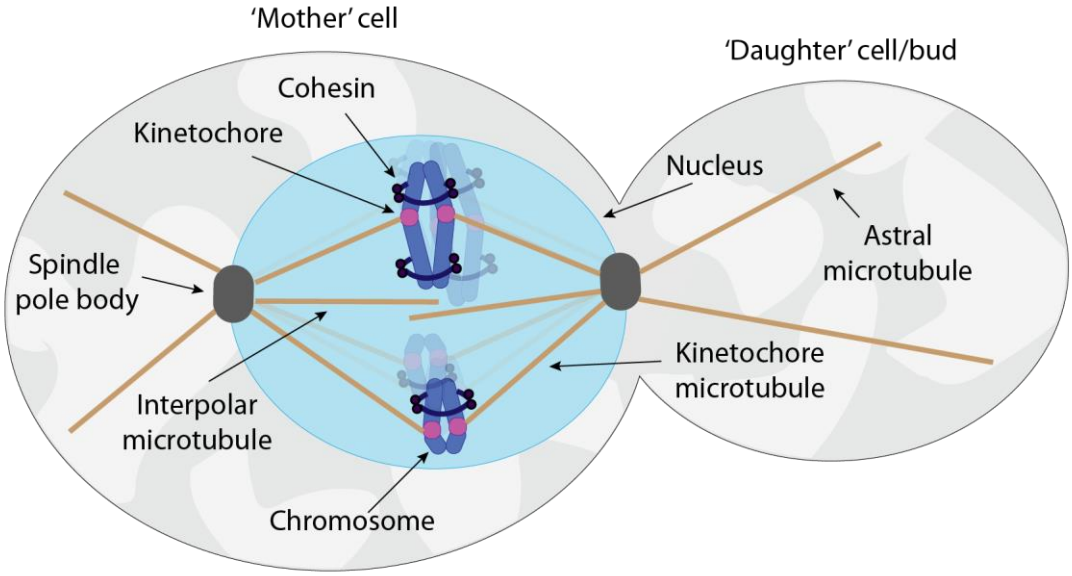

Figure 2: Chromosome segregation of budding yeast

Schematic representation of chromosome segregation: the newly replicated sister chromatids (chromosomes), held together by cohesin, are pulled apart into the two emerging daughter cells. Budding yeast undergoes a so-called 'closed' mitosis in which the nuclear membrane stays intact during the whole process. Microtubules are organised from the two spindle pole bodies, which are embedded in the nuclear envelope at the opposite poles. One can differentiate between microtubules attaching to the kinetochores and therefore to the centromeres (kinetochore microtubules), those extending towards the opposite spindle pole body (interpolar microtubules) and those extending towards the cell membrane (astral microtubules). Whilst the latter are important throughout the whole cell cycle to position the nucleus, interpolar microtubules are essential for spindle stability during mitosis. The kinetochores are responsible for sensing if all sister chromatids are correctly attached, a process involving establishment of tension.

The overall architecture and mode of function is conserved between budding yeast and man, as are the majority of proteins involved (Meraldi et al., 2006). The budding yeast kinetochore, however, represents a much more simplified version and therefore has been used by many research groups as a model to study kinetochore function. Yet even this 'simple' budding yeast kinetochore consist of more than 60 different proteins (Santaguida and Musacchio, 2009, McAinsh et al., 2003), which are orchestrated by a variety of regulative factors to facilitate faithful segregation. Defects in the kinetochore or its regulation result in chromosome mis-segregation, which can lead to cell cycle arrest and apoptosis or to chromosomal instability (Schuyler et al., 2012). The latter is observed in many types of cancer (Tanaka and Hirota, 2009, Jallepalli and Lengauer, 2001), which underlines the importance of understanding chromosome segregation and its regulation, thus enabling a better understanding of how it malfunctions in cancer.
As this thesis is describing work on the essential budding yeast kinetochore complex CBF3, the following chapters are focussing on the architecture, function and regulation of the budding yeast kinetochore and centromere, whilst trying to emphasise similarities, but also discrepancies towards higher eukaryotes.

1.2 The Centromere

Whilst the overall structure of the kinetochore and the majority of its proteins are conserved (Meraldi et al., 2006), the centromeric sequence is highly divergent (Steiner and Henikoff, 2015). Indeed, centromeres can be separated into two groups, the point centromeres of budding yeasts and the regional centromeres of higher eukaryotes (Figure 3). The former are about 125bp long and consist of three elements, CDEI, II and III. Whilst CDEII is exceptionally A-T rich but vary in sequence and length, CDEI and III are conserved (Jehn, 1991, Clarke and Carbon, 1980) and are recognised and bound in a sequence-specific matter by Cbf1 and CBF3 respectively (Cai and Davis, 1989, Lechner, 1991, Ng and Carbon, 1987). In contrast, the regional centromeres of higher eukaryotes are much longer, varying in size from a few thousand to mega base-pairs and their sequence is not conserved but consists of repetitive elements (Malik and Henikoff, 2002).

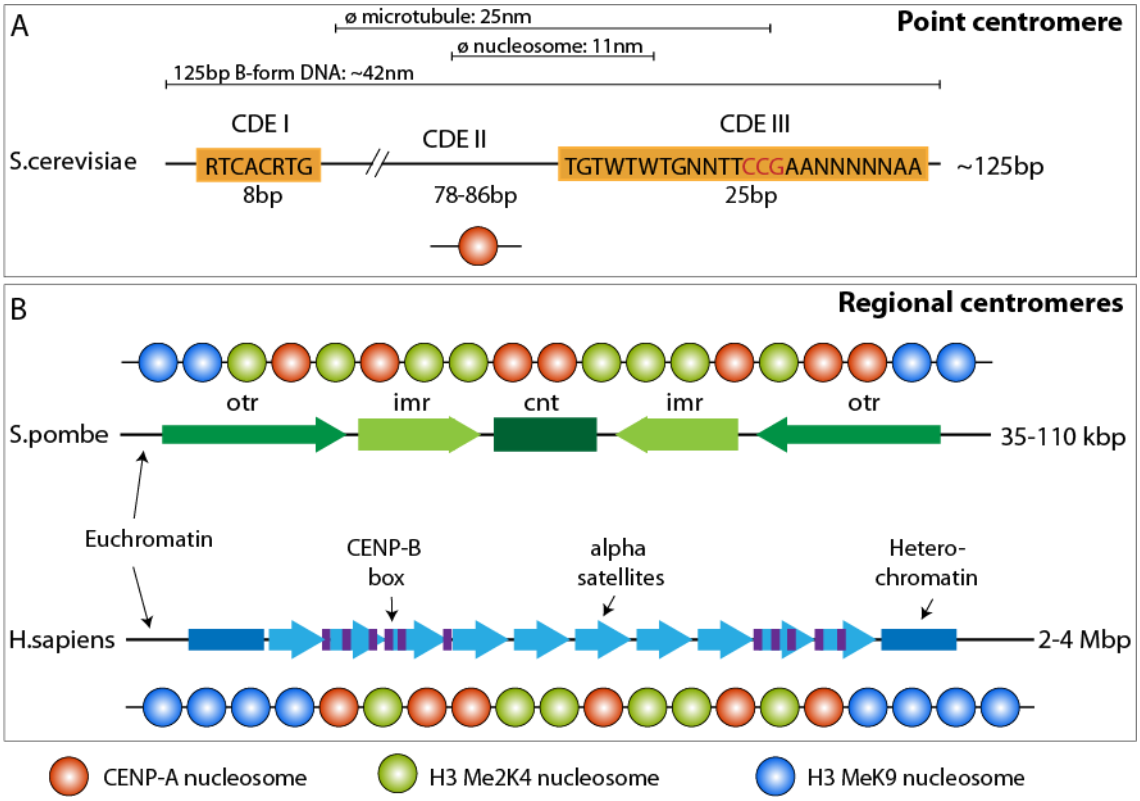

Figure 3: Point vs. regional centromeres

A The short (~125bp) point centromere of budding yeast consists of three elements, CDEI, II and III. CDEII is non-conserved, but AT-rich and varies in length, whereas the other two elements are conserved sites for specific protein interaction. The essential CCG triplet in the CDEIII is highlighted in red. One centromeric nucleosome binds the entire centromeric sequence. **B** The long regional centromere of *S.pombe* consists of a central core (cnt), flanked by inner repeats (imr) and outer repeats (otr). The latter is made out of three elements: the dg, cen253 and dh repeats. Dg repeats are most conserved between the three chromosomes of *S.pombe* and are essential for centromere function alongside the cnt region. Centromeric nucleosomes, as well as H3 nucleosomes, featuring lysine-24 double methylation, as well as lysine-9 methylation in neighbouring heterochromatin. Human centromeres are up to Mbp long and consist mainly of alpha satellite DNA, of various sequence conservation and other elements inserted, such as the CENP-B box bound by CENP-B.

Beside these differences all centromeres are epigenetically defined by the occurrence of a centromere-specific nucleosome, in which the canonical histone H3 is replaced by the centromere-specific variant CENP-A (Cse4 in *S.cerevisiae*; Palmer et al., 1987, Stoler et al., 1995, Allshire and Karpen, 2008). In point centromere species, it is believed that a single CENP-A nucleosome binds the entire

centromere and is connected through the kinetochore to a single microtubule (Meluh et al., 1998, Furuyama and Biggins, 2007, Camahort et al., 2009, Cole et al., 2011). On the other hand, in regional centromere species many copies of CENP-A nucleosomes are bound to the long centromeric sequence, which in turn are connected to many spindle microtubules (Zinkowski et al., 1991, Bodor et al., 2014). The consequently larger kinetochore is believed to be assembled from a repetition of the basic 'budding yeast kinetochore unit'; a widely accepted view, which is also known as the 'repeat subunit' model (Zinkowski et al., 1991, Joglekar et al., 2008, Santaguida and Musacchio, 2009, Burrack and Berman, 2012). These findings underline the importance of the simpler budding yeast kinetochore as the choice model organism, despite its different centromere structure.

1.2.1 Centromere-specific nucleosome

As mentioned above, a unique feature of both point and regional centromeres is the occurrence of the centromere-specific nucleosome. Whilst it is generally accepted that the centromeric nucleosome is defined by the replacement of H3 with CENP-A, its composition and structure is still a matter of debate. Whilst *in vitro* reconstituted centromeric nucleosomes tend to form conventional octamers, as determined by biochemistry (Yoda et al., 2000, Camahort et al., 2009, Kingston et al., 2011) and the crystal structure of the human CENP-A nucleosome (Tachiwana et al., 2011), there are several *in vivo* evidences of alternative structures. One of these is the so-called 'hemisome', consisting of only one copy of CENP-A, H4, H2A and H2B. It was first purified from interphase *Drosophila melanogaster* cells after crosslinking and immunoprecipitation of CENP-A^{CID} (Dalal, 2007) and could be stably reconstituted on 78bp AT-rich DNA with budding yeast proteins (Furuyama et al., 2013). Beside this hemisome structure, there is evidence in *S.cerevisiae* that the non-histone protein Scm3 can replace histones H2A and H2B from pre-assembled octameric nucleosomes (Mizuguchi et al., 2007). That and results from CHIP assays, which showed that H2A and H2B is absent at centromeric chromatin, led to the suggestion of a hexameric Cse4/H4/Scm3 structure (Mizuguchi et al., 2007). Furthermore, *in vivo* DNA topology studies showed that functional budding yeast centromeres introduce positive supercoiling (Diaz-Ingelmo et al., 2015, Furuyama and Henikoff, 2009). A canonical nucleosome, as well as the CENP-A nucleosome crystal

structure, however, wraps DNA in a right-handed way, which subsequently leads to negative supercoiling (Figure 4).

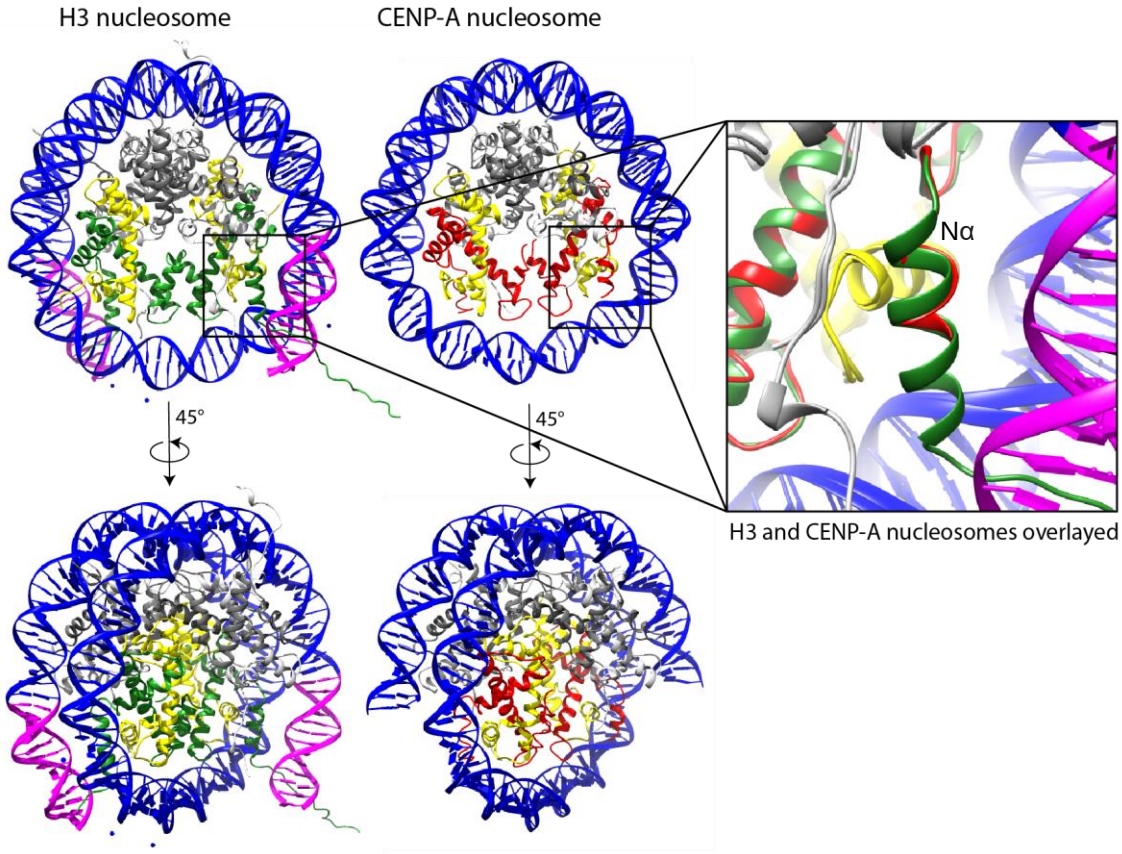

Figure 4: H3 vs. CENP-A nucleosome

Side-to-side comparison of the H3 (left; 1AOI) and CENP-A (right; 3AN2) nucleosome crystal structures in two orientations. Whereas the histones themselves pack in a very similar fashion, the main difference between the two structures are the DNA ends: whereas all 146bp of DNA are tightly wrapped around the H3-containing histone octamer, 13 base pairs from both ends of the DNA in the CENP-A crystal structure are not resolved, indicating a loose interaction (the corresponding base pairs in the H3 nucleosome structure are highlighted in pink). One should note that both H3 as well as CENP-A nucleosomes were reconstituted with 146bp DNA. The boxed close-up view shows the difference in the Nα helix between H3 (green) and CENP-A (red). The latter is shorter and its preceding tail completely disordered, explaining the looser DNA contacts in this region. All other differences are reviewed in detail in Tachiwana et al., 2011. Histones are coloured as following: H2A… light grey; H2B… dark grey; H4… yellow; H3… green; CENP-A… red.

On the other hand, many features of the CENP-A nucleosome crystal structure align with biochemical data (Tachiwana et al., 2011, Moreno-Moreno et al., 2017). One of these features is, that the CENP-A nucleosome binds DNA ends less stably, resulting

in a more open conformation as is seen for the H3 nucleosomes (Figure 4). This explains biochemical data from human, as well as budding yeast and drosophila, showing that their centromeric nucleosomes protect less DNA from nuclease treatment than a conventional nucleosome (Yoda et al., 2000, Cole et al., 2011, Kingston et al., 2011). Two arguments speak for a conserved structure of the centromeric nucleosome between budding yeast and human. First, the high sequence conservation of the proteins in their histone-fold-domain (HFD; Sullivan et al., 1994, Stoler et al., 1995), and second the fact that overexpression of Cse4 can rescue a RNAi induced CENP-A deletion in human cells (Wieland et al., 2004). To clarify the composition of the budding yeast centromeric nucleosome, it will be necessary to investigate the structure of the yeast centromere bound by the nucleosome and other centromere-associated proteins, like the essential, four-subunit centromere-binding factor 3 (CBF3). A limiting factor for the reconstitution of such a bigger assembly so far, was the inability to recombinantly express and purify the budding yeast CBF3 complex in useful quantities.

Besides the composition of the centromeric nucleosome, another interesting question in the field is how is CENP-A specifically recruited to the centromere. Besides the differences in centromere sequence, a conserved histone chaperone, HJURP in human and Scm3 in yeasts, is involved in this process (Camahort et al., 2007, Shuaib et al., 2010). It forms a stable and soluble complex with CENP-A and H4 and is involved in their assembly into centromeric chromatin (Shuaib et al., 2010, Dechassa et al., 2014). A variety of proteins act to prime centromeric sequence for CENP-A incorporation through their interaction with $HJURP^{Scm3}$ in regional centromere species (Hayashi et al., 2004, Fujita et al., 2007, Nardi et al., 2016, An et al., 2018; reviewed in Musacchio and Desai, 2017). In budding yeast, however, it is suggested, that Scm3 is recruited to the centromere through its interaction with the centromere binding complex CBF3 (Camahort et al., 2007, Cho and Harrison, 2012). The latter is discussed in more detail in 1.4.3.

1.3 The Kinetochore

The main function of the kinetochore is to connect the centromeres of sister chromatids to spindle microtubules of the opposite pole. This is facilitated through one of the biggest macromolecular assemblies in the cell. One can divide the overall

kinetochore structure in three different layers: The inner layer which recognises and binds the centromere; the outer layer which interfaces with spindle microtubules and the middle layer which acts as a bridge between the other two and is involved in the regulation of function (reviewed in Santaguida and Musacchio, 2009 and Musacchio and Desai, 2017). Beside the physical anchoring of centromeres to spindle microtubules, the kinetochore is also responsible for the verification of correct attachment, the activation of the spindle assembly checkpoint (SAC) and also participates in microtubule dependent force generation for movements of chromosomes (Joglekar et al., 2010, Suzuki et al., 2016). The budding yeast kinetochore consists of more than 60 different proteins, many of which cluster in subcomplexes and most are conserved to humans (Figure 5, Table 1; McAinsh et al., 2003, De Wulf et al., 2003, Meraldi et al., 2006). Although featuring the simplest kinetochore, much of the molecular architecture and function is still unknown.

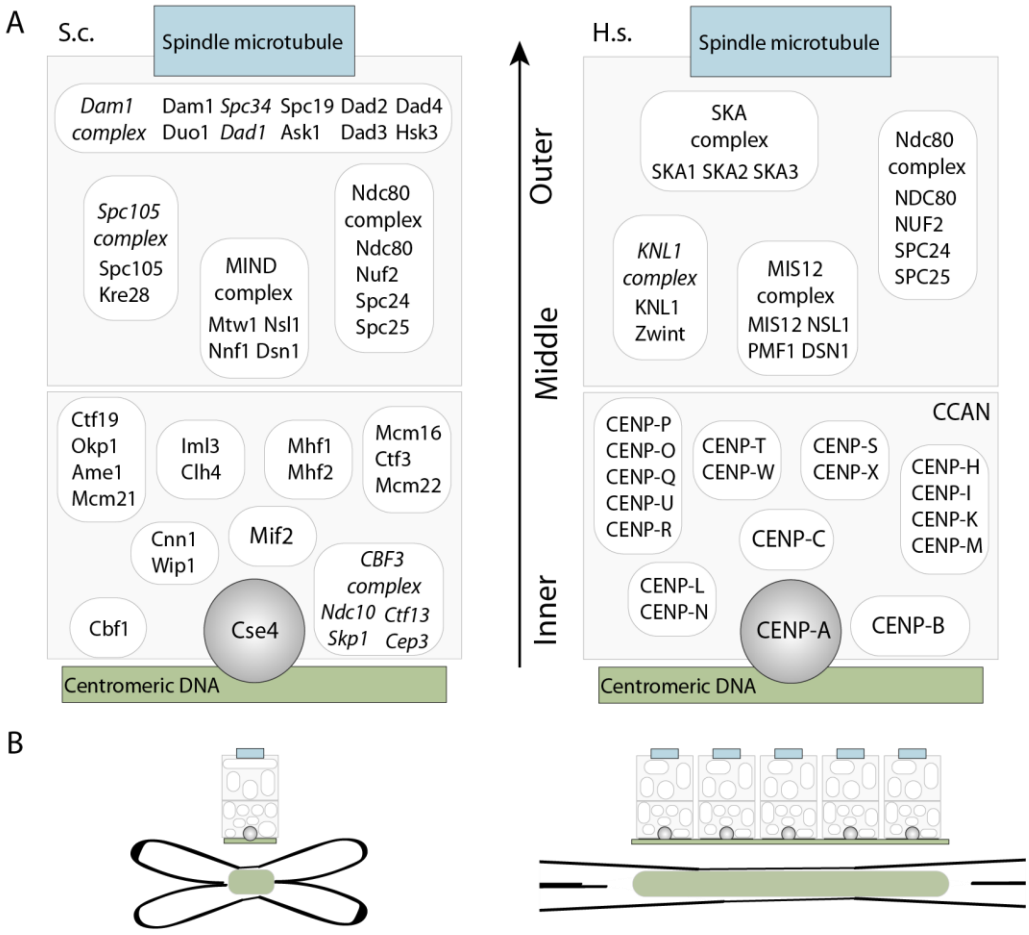

Figure 5: Kinetochore schematic

A Illustration of kinetochore sub-complexes from budding yeast (left) and human (right), excluding regulator proteins like the chromosomal passenger complex and SAC

components. Some available structures are shown in Figure 6 and all kinetochore proteins are listed in Table 1. **B** Schematic of the 'repeat subunit' model, demonstrating that the human kinetochore consists out of many repeats of the basic budding yeast kinetochore unit. Therefore, each of the CENP-A nucleosomes is connected to a spindle microtubule, whereas in budding yeast the single Cse4 nucleosome connects to only one spindle microtubule.

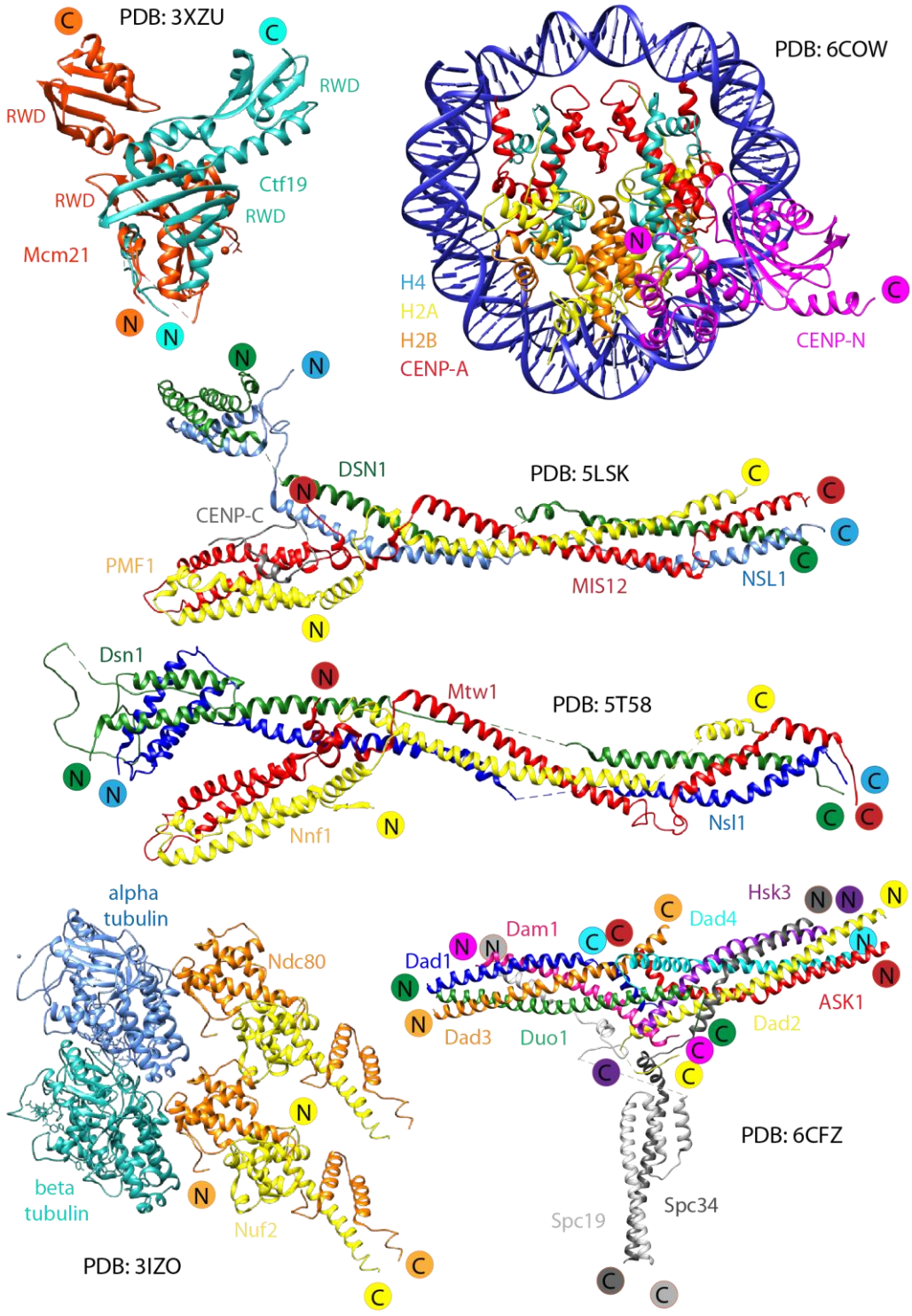

Figure 6: Selection of structures of kinetochore subunits

3XZU: crystal structure of the heterodimer of the *K.lactis* Mcm21 and Ctf19, part of the four subunit COMA complex, featuring four RWD domains. **6COW**: cryoEM structure of the CENP-A nucleosome bound by CENP-N, which interacts with CENP-A through its L1 loop. **5LSK**: Crystal structure of the budding yeast MIND complex, made out of four subunits and participating an elongated structure which connects both to the inner kinetochore through its interaction with CENP-CMif2 and to the outer kinetochore complexes like Ndc80 complex. **5T58**: crystal structure of the human homologue, the MIS12 complex, exhibiting a similar structure. Here also a small part of CENP-C could be co-crystallised and resolved. **3IZO**: cryoEM structure of Ndc80/Nuf2 bound to microtubule, shown as only one alpha and beta tubulin. **6CFZ**: CryoEM analysis of the DASH/Dam1 ring whereas only one heterodecameric unit of the ring is shown.

Below, an overview of the current understanding of the budding yeast kinetochore architecture is given, starting from the inner layer of centromere associated proteins and moving towards the outer layer of microtubule binding subunits. Differences and similarities to higher eukaryotes with regional centromeres are discussed. All main kinetochore proteins and their conservation between the three largely studied organism are listed in Table 1. Finally, the regulative functionalities of the kinetochore are described, including the error-correction and the spindle assembly checkpoint (SAC).

Table 1: Overview list of main kinetochore proteins

Kinetochore proteins in budding yeast, fission yeast and humans grouped by subcomplexes. It should be noted that this list is not comprehensive and additional kinetochore proteins exist. For a more elaborate list see Musacchio and Desai, 2017.

Kinetochore proteins			**Subcomplexes**		
S.cerevisiae	***S.pombe***	***H.sapiens***	***S.cerevisiae***	***S.pombe***	***H.sapiens***
Ndc10	-	-	CBF3	-	-
Cep3	-	-		-	-
Ctf13	-	-		-	-
Skp1	-	-		-	-
Cse4	Cnp1	CENP-A	-	-	-
Cbf1	-	-	-	-	-
Mif2	Cnp3	CENP-C	-	Part of CCAN	Part of CCAN
-	-	CENP-B			
Cnn1	-	CENP-T	Ctf19 complex	Sim4 complex	CCAN
Wip1	-	CENP-W			

Mcm16	Fta3	CENP-H			
Ctf3	Mis6	CENP-I			
-	Sim4	CENP-K			
-	-	CENP-M			
Mcm21	Mal2	CENP-O			
Ctf19	Fta2	CENP-P			
Okp1	-	CENP-Q			
-	-	CENP-R			
Ame1	Fta4	CENP-U			
Iml3	Fta1	CENP-L			
Chl4	Mis15	CENP-N			
Mhf1	-	CENP-S			
Mhf2		CENP-X			
Mcm22	-	-			
-	Mis17	-			
-	Fta5	-			
-	Fta6	-			
-	Fta7	-			
Spc105	Spc7	KNL1	Spc105 complex	Spc7 complex	KNL-1 complex
Kre28	-	Zwint			
Mtw1	Mis12	MIS12	MIND complex	Mis12 complex	MIS12 complex
Dsn1	Mis13	DSN1			
Nnf1	Nnf1	PMF1			
Nsl1	Mis14	NSL1			
Ndc80	Ndc80	NDC80	Ndc80 complex	Ndc80 complex	Ndc80 complex
Nuf2	Nuf2	NUF2			
Spc24	Spc24	SPC24			
Spc25	Spc25	SPC25			
Dam1	Dam1	-	Dam1 complex	Dam1 complex	-
Duo1	Duo1	-			
Spc34	Spc34	-			
Dad1	Dad1	-			
Spc19	Spc19	-			
Ask1	Ask1	-			
Dad2	Dad2	-			
Dad3	Dad3	-			
Dad4	Dad4	-			
Hsk3	Hsk3	-			
-	-	ROD	-	-	RZZ complex
-	-	ZW10			
-	-	Zwilch			
-	-	SKA1	-	-	Ska complex
-	-	SKA2			
-	-	SKA3			
Mad2	Mad2	MAD2	Mitotic checkpoint complex	Mitotic checkpoint complex	Mitotic checkpoint complex
Bub3	Bub3	BUB3			
Mad3	Mad3	BUBR1			
Cdc20	Slp1	CDC20			
Ipl1	Ark1	Aurora B	Chromosome Passenger complex	Chromosome Passenger complex	Chromosome Passenger complex
Sli15	Plc1	INCENP			
Bir1	Cut7	Survivin			
Nbl1p	-	Borealin			

1.3.1 Inner kinetochore

One of the first budding yeast kinetochore complexes to be described was the CBF3 complex, which binds sequence-specifically to the centromere and is required for the recruitment of all other kinetochore proteins to the centromeric DNA (Ng and Carbon, 1987, Lechner, 1991, Goh, 1993, Strunnikov et al., 1995, Connelly and Hieter, 1996). CBF3 consists of four subunits, all of which are essential and mutations of these lead to a complete loss of kinetochore function *in vivo*, as well as a loss of centromere/spindle microtubule attachment *in vitro* (Goh, 1993, Sorger, 1994, He et al., 2001).

Other inner kinetochore proteins bound directly to the centromere include the homodimeric protein Cbf1 (Bram and Kornberg, 1987), the centromere-specific Cse4 nucleosome (Meluh et al., 1998) and Mif2 (Meluh and Koshland, 1997). Cbf1 recognises and specifically binds the centromere element CDEI and although it is not essential, its deletion leads to an increased loss of a mitotic CEN plasmid (Cai and Davis, 1989, Mellor et al., 1990). On the other hand, the centromeric nucleosome and Mif2 are absolutely essential for kinetochore establishment and function (Brown et al., 1993). The latter has been shown to specifically bind to the Cse4-nucleosome (Falk et al., 2015, Xiao et al., 2017), as well as other inner and middle kinetochore complexes (see 1.3.2). How all of these proteins can bind such a short stretch of DNA is still debated, as is the conformation of the centromeric nucleosome (see 1.2.1). Structural studies of the whole inner kinetochore will be crucial to clarify these issues. Whilst the centromere-specific nucleosome and Mif2 (CENP-C in human) are conserved between budding yeast and regional centromere species, Cbf1 and CBF3 are not (Cai and Davis, 1989, Mellor et al., 1990, Meluh and Koshland, 1995). However, additional proteins are found in higher eukaryotic centromeres, such as CENP-B and members of the CCAN (constitutive centromere-associated network; reviewed in Musacchio and Desai, 2017).

CENP-B was one of the earliest centromere binding proteins identified, however, its role is still unclear and it is not essential for kinetochore function (Earnshaw and Rothfield, 1985). It binds, however, to the 17bp CENP-B boxes of the human centromere (K-type repeats in *S.pombe*) and introduces DNA bending (Tanaka et al., 2001). Interestingly both Cbf1 and CBF3 also influence DNA conformation

through bending and possibly looping of the DNA (Niedenthal et al., 1993, Pietrasanta, 1999).

The budding yeast homologue to the CCAN network is the Ctf19 complex, with many proteins being conserved (Table 1). The CENP-TWXS complex of the CCAN directly interacts with the long regional centromere (Hori et al., 2008, Nishino et al., 2012) with the homologues yeast proteins likely playing a similar role (Schleiffer et al., 2012, Pekgoz Altunkaya et al., 2016, Singh et al., 2010). The centromere localisation depends on CENP-A^{Cse4}, but no direct interaction exists with the CENP-A^{Cse4} nucleosome (Hori et al., 2008). Besides these DNA-binding subunits, the CCAN network consists of eleven other proteins, which themselves are organised into individual subcomplexes (reviewed in Musacchio and Desai, 2017). It is involved in the recruitment of CENP-A (Carroll et al., 2009, Suma et al., 2018), SAC proteins (Liu et al., 2003) and cohesin (Lacefield, 2017), as well as in establishing the connection between the centromere and outer microtubule-binding kinetochore proteins (Hori et al., 2008, Huis In 't Veld et al., 2016, Schleiffer et al., 2012). The first proteins of the Ctf19 complex to be described were members of the COMA complex, Ctf19, Okp1 and Mcm21. This sub-complex has been found to associate with the budding yeast centromere *in vivo* through its interaction with CBF3, Cse4 as well as Mif2 (Ortiz et al., 1999). Other subunits were found by genetic lethality screens, affinity purification and mass spectrometry and a variety of subcomplexes were successfully reconstituted (Hyland et al., 1999, Measday et al., 2002, De Wulf et al., 2003, Hornung et al., 2011, Hornung et al., 2014, Schmitzberger and Harrison, 2012, Schmitzberger et al., 2017, Weir et al., 2016). Most of the Ctf19 complex subunits are conserved to the CCAN of higher eukaryotes and they also share most of their functionality. One known exception is the involvement of the CCAN in CENP-A loading in higher eukaryotes, as all known Ctf19 subunits are dependent on prior CBF3, Mif2 and Cse4 localisation to the centromere (Ortiz et al., 1999, Hyland et al., 1999, De Wulf et al., 2003).

1.3.2 Middle kinetochore

Two complexes, MIND and Spc105, form the middle part of the budding yeast kinetochores. Both complexes are conserved towards higher eukaryotes (see Table 1) and are involved in the regulation of kinetochore function, as well as forming the

bridge between the above described centromere associated subcomplexes and the microtubule binding subunits (Euskirchen, 2002). Together with the outer microtubule-binding subcomplex Ndc80 (see 1.3.3) they can also be classified as the KMN network, named after the human homologues KNL1, MIS12 and Ndc80 complexes (Cheeseman et al., 2006).

The MIND complex consists of four proteins, Mtw1, Nnf1, Nsl1 and Dsn1, and exhibits an elongated structure with the C- and N-termini of all four subunits clustering on both ends (Maskell et al., 2010, Dimitrova et al., 2016, Petrovic et al., 2016). A region near the C-termini of Dsn1 and Nsl1 binds to the Spc24/25 subunits of the Ndc80 complex and therefore establishes connection to the outer microtubule-binding site of the kinetochore. The opposite end of the structure connects to the centromere-associated inner kinetochore through a conserved interaction with Mif2 (Hornung et al., 2011, Dimitrova et al., 2016, Petrovic et al., 2016). This connection is additionally stabilised through the COMA complex, although this has only been shown for budding yeast (Hornung et al., 2014). In both human and budding yeast, however, phosphorylation of Dsn1 by Ipl1 kinase (Aurora B in human) reliefs an auto-inhibitory mechanism and therefore increases the binding affinity between the Mif2 and the MIND complex significantly (Akiyoshi et al., 2013, Yang et al., 2008, Welburn et al., 2010, Dimitrova et al., 2016, Petrovic et al., 2016). Besides bridging the inner and outer kinetochore, the MIND complex also binds the Spc105 complex, and therefore is often described as the platform for KMN assembly (Petrovic et al., 2010, Petrovic et al., 2014). Both in human, as well as budding yeast there is an additional linking between inner and outer kinetochore proteins through a direct connection between the CENP-T^{Cnn1} subunits of the CCAN^{Ctf19} and the Ndc80 complex (Malvezzi et al., 2013, Nishino et al., 2013, Pekgoz Altunkaya et al., 2016, Huis In 't Veld et al., 2016). This additional bridging is supposed to help to recruit additional Ndc80 complexes to the kinetochore and increase binding stability of the spindle microtubules (reviewed in Musacchio and Desai, 2017).

The Spc105 complex of the middle kinetochore, on the other hand, does not contribute to inner and outer kinetochore cohesion, but builds a platform for the assembly of SAC components. It consists out of two conserved subunits, Spc105 and Kre28 (see Table 1) and it connects to the MIND complex through a C-terminal region of Spc105 (Petrovic et al., 2010, Petrovic et al., 2014). The large subunit Spc105, or KNL1 in human, recruits a number of proteins including the SAC proteins

Bub1 and Bub3 through its large intrinsically disordered N-terminal part (Liu et al., 2010, Rosenberg et al., 2011, Primorac et al., 2013).

1.3.3 Outer kinetochore

The outer kinetochore's function is to establish an end-on connection to the spindle microtubules. This connection ensures the separation of the sister chromatids by utilising the force generated by depolymerisation of the microtubules.

The budding yeast outer kinetochore is mainly formed by two complexes, the Dam1 and the Ndc80 complex. Only the latter, however, is conserved to human (Zheng et al., 1999, Cheeseman et al., 2006). It consists out of four subunits which together form a long coiled-coil with globular domains on each end (Wei et al., 2007, Ciferri et al., 2008). These globular domains bind to microtubules on one end and to the MIND complex on the other (Wei et al., 2005, Ciferri et al., 2008). The Ndc80 complex is one of the best structurally and functionally described subcomplexes of the kinetochore and binds microtubules through two calponin-homology (CH) domains of Ndc80 and Nuf2 subunits, as well as through the structurally disordered N-terminal tail of Ndc80 (reviewed in Musacchio and Desai, 2017, Ciferri et al., 2008, Alushin et al., 2012). The latter is also involved in the regulation of microtubule binding through Aurora B^{Ipl1} kinase (Alushin et al., 2012), which is discussed in more detail in 1.3.4. Multiple copies of the Ndc80 complexes exist in parallel and might bind to the same microtubule in a cooperative manner (Alushin et al., 2010, Suzuki et al., 2016).

The second main player in budding yeast is the Dam1 complex (Hofmann et al., 1998), which is responsible for microtubule/kinetochore coupling and is able to track growing and shrinking microtubules (Asbury et al., 2006). Although Dam1 proteins are not conserved, a functional homologue, the SKA complex, exists in human (Hanisch et al., 2006, Gaitanos et al., 2009). Both interact with the Ndc80 complex and enhance the microtubule attachment helping to form load-bearing connection (Tien et al., 2010, Lampert et al., 2010, Schmidt et al., 2012). Furthermore, both are negatively regulated by Aurora B kinase, as phosphorylation of Dam1 and Ska complex subunits decreases their binding affinity for the Ndc80 complex and microtubules respectively (Tien et al., 2010, Schmidt et al., 2012). This regulation is part of the error correction mechanism of kinetochores, discussed in the following section.

1.3.4 Error correction

To ensure proper chromosome segregation the cell not only needs to establish a connection between centromere and spindle microtubules, it also needs to sense if this attachment is correct (reviewed in Lampson and Grishchuk, 2017). In budding yeast two possible connections can be made (Figure 7). The first and correct one is called amphitelic and describes an attachment of the two sister chromatids to spindle microtubules of the opposite pole. The second possibility is an attachment of the two sister chromatids to spindle microtubules of the same pole, which is called syntelic. This erroneous form of attachment needs to be corrected to ensure proper chromosome segregation. A third form of incorrect connection can occur in regional centromere species, which have multiple spindle microtubules connected to one sister chromatid (Zinkowski et al., 1991, Nicklas, 1997, Bodor et al., 2014). This gives rise to the possibility of a merotelic attachment, where one or both sister chromatids are connected to spindle microtubules from both poles, rather than from one. The mechanism to ensure proper attachment is believed to be conserved, however, and involves a tension-regulated stabilisation of amphitelic connections (Nicklas and Koch, 1969, Cane et al., 2013). Only these are slightly pulled apart by microtubules and therefore build up a tension. Syntelic and merotelic attachments, on the other hand, fail to do so and are therefore resolved. A new spindle microtubule can then be bound and this attachment is tested again by the same mechanism. Eventually the correct form of attachment is made (Nicklas and Ward, 1994).

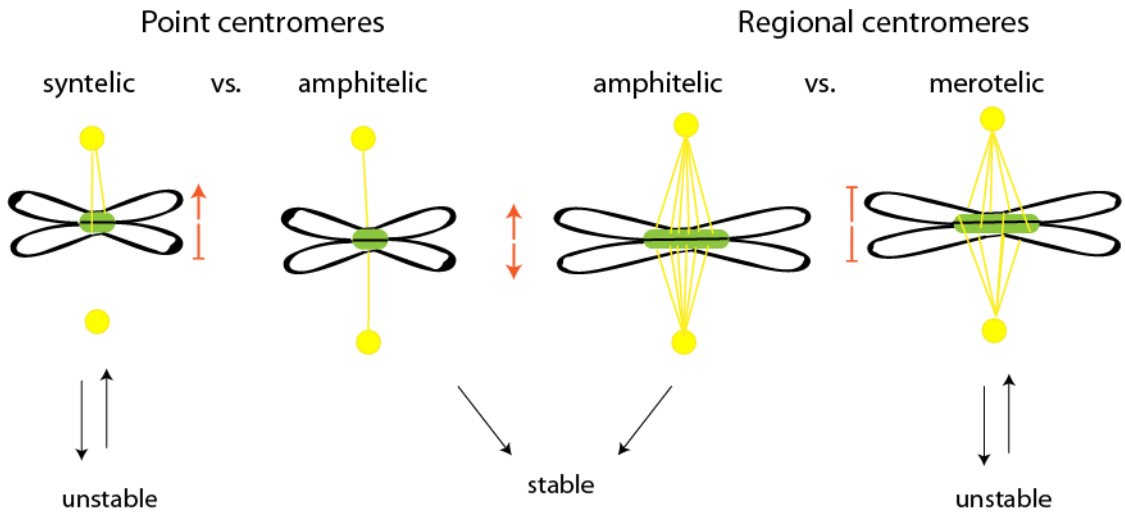

Figure 7: Microtubule/kinetochore attachments during mitosis

Two forms of attachments are possible in point centromere species: Syntelic attachments, both sister chromatids are connected to microtubule of one spindle pole, do not build tension and are corrected, whereas the amphitelic attachment is stabilised. In regional centromeres multiple spindle microtubules are connected to the sisters, enabling also a merotelic attachment, where one or both sister is connected to spindle microtubules of both poles.

Whilst this is the generally accepted model of tension-regulated error correction, there are still many unanswered aspects, especially concerning the more complicated scenario of a merotelic attachment (reviewed in Lampson and Grishchuk, 2017). In addition to this tension-regulated mechanism, other factors have been shown to influence error correction, including the back-to-back geometry of the sister kinetochores (Östergren, 1951, Nicklas and Ward, 1994, Anderson, 2009).

The molecular mechanism of error correction involves the conserved chromosome passenger complex (CPC), consisting of four subunits including the Aurora B^{Ipl1} kinase, INCENP^{Sli15}, Survivin^{Bir1} and Borealin^{Nbl1p} (see Table 1; Kaitna et al., 2000, Murata-Hori and Wang, 2002, Carmena et al., 2009). This complex acts at different locations at different times during mitosis and coordinates not only correction of chromosome-microtubule attachment errors, but also the activation of the SAC, establishment of a bipolar spindle and cytokinesis (reviewed in Carmena et al., 2009). For its function in error correction, however, it localises to the inner kinetochore (Kawashima et al., 2010, Wang et al., 2010, Yoon and Carbon, 1999), whilst Aurora B targets include the outer kinetochore proteins Ndc80 and KNL1, as well as subunits of the Dam1 or Ska complex (Welburn et al., 2010, Keating et al., 2009, Chan et al., 2012). Phosphorylation of these substrates reduces microtubule binding affinity and promotes catastrophic depolymerisation of the attached microtubules (Sarangapani et al., 2013, Umbreit et al., 2012, DeLuca and Musacchio, 2012, Zaytsev et al., 2014, Zaytsev et al., 2015). The current model is that Aurora B can only access its substrates when the kinetochore/microtubule attachment is wrong, and therefore not under tension. Once tension is applied Aurora B is spatially separated from its targets and unable to phosphorylate them (Liu et al., 2009, Keating et al., 2009, Welburn et al., 2010, Krenn and Musacchio, 2015). Studies which support this model have shown that perturbing centromere localisation of the CPC disrupts error correction (Wang et al., 2010, Kelly et al., 2010) and

artificial targeting of Aurora B to the outer kinetochore decreases microtubule attachment (Liu et al., 2009). For stabilisation of microtubule attachment Aurora B substrates are dephosphorylated by the opposing protein phosphatase 1 (PP1; Francisco et al., 1994, Emanuele et al., 2008). PP1 is specifically recruited to the outer kinetochore through interaction with KNL1. In the presence of Aurora B, however, this binding site is phosphorylated, which disrupts PP1 localisation. Stabilisation through PP1 can therefore only occur in the absence of Aurora B (Liu et al., 2010).

1.3.5 Spindle assembly checkpoint

Besides above discussed error correction, the kinetochore also couples successful sister chromatid bi-orientation with the mitotic exit, a process called the spindle assembly checkpoint (SAC; reviewed in Joglekar, 2016). The SAC ensures that the cell only progresses with mitosis and subsequent cytokinesis when all sister chromatids are properly attached to spindle microtubules. This is facilitated by the recruitment of SAC proteins to unattached kinetochores, which generates the so-called ‘wait-anaphase’ signal (McIntosh, 1991, Bernat et al., 1991, Spencer and Hieter, 1992). The progression from anaphase onwards is dependent on the anaphase promoting complex/cyclosome C (APC/C), which is activated by its cofactors Cdc20 and Cdh1 (Sudakin et al., 1995, Irniger et al., 1995, Schwab et al., 1997, Hwang et al., 1998). Unattached kinetochores recruit various SAC proteins, which leads to the formation of the mitotic checkpoint complex (MCC) consisting out of four proteins, Bub3, Mad2, Mad3 and Cdc20 (see Table 1; Hardwick et al., 2000, Sudakin et al., 2001). This complex formation, therefore, depletes free cytosolic Cdc20 and also inactivates Cdc20 already bound by the APC/C, which subsequently leads to complete inactivation of APC/C (Nilsson et al., 2008, Izawa and Pines, 2015). Besides inhibiting anaphase onset, the SAC also protects cohesion between the sister chromatids to inhibit segregation before proper attachment of all sister chromatids are made (Yamamoto et al., 1996, Musacchio, 2015, Lara-Gonzalez et al., 2012, Chen and Liu, 2014).

The kinetochore proteins involved in the SAC component recruitment are Ndc80 and Spc105, as well as the kinase Mps1 (Kemmler et al., 2009, Hiruma et al., 2015, Ji et al., 2015, Dou et al., 2015). The latter binds to the CH domains of Ndc80, which are

also involved in Ndc80/microtubule interaction (DeLuca et al., 2003, DeLuca et al., 2006). This competitive binding therefore favours a recruitment to unattached kinetochores. Bound Mps1 phosphorylates Spc105 at its MELT repeats, which leads to the recruitment of SAC proteins Bub3-Bub1, and subsequently Mad3, Cdc20, Mad1-Mad2 (Kemmler et al., 2009, Hiruma et al., 2015, Ji et al., 2015, Dou et al., 2015, Dimitrova et al., 2016, Jia et al., 2016) and the formation of the MCC complex. In addition, Aurora B has been shown to promote Mps1 binding to Ndc80, coupling its error correction role to promoting SAC signalling on unattached or wrongly attached kinetochores (Krenn and Musacchio, 2015).

Multiple mechanisms seem to disrupt SAC signalling once kinetochores reach a proper attachment. Firstly, upon end-on microtubule attachment Mps1 binding is disrupted, reducing the amount of bound Mps1 (Joglekar, 2016). Secondly, end-on attachment also leads to a different conformation of Ndc80 and Spc105, which spatially separates the two. Mps1, bound by Ndc80, therefore cannot phosphorylate Spc105 anymore (Wang et al., 2008, Joglekar et al., 2009, Wan et al., 2009, Schittenhelm et al., 2009, Aravamudhan et al., 2015), leading to dephosphorylation by PP1 bound directly by Spc105 (London et al., 2012). Additionally, other proteins have been indicating to play a role in SAC silencing, explaining the immediate response once kinetochores are properly attached (Espert et al., 2014, Sivakumar et al., 2016, Howell et al., 2001).

1.4 The CBF3 complex and kinetochore establishment

In budding yeast kinetochore establishment on the centromere starts with the specific binding of CBF3 (reviewed in McAinsh et al., 2003). Defects in CBF3 abolish kinetochore function (Goh, 1993, Sorger, 1994, He et al., 2001) and all other known kinetochore proteins, with the exception of Cbf1, are dependent on CBF3 for their recruitment (Sorger, 1994, He et al., 2001, Enquist-Newman et al., 2001, Janke et al., 2001, Jones et al., 2001). Despite this, CBF3 only directly interacts with Cbf1 and the COMA subunits of the Ctf19 complex (Ortiz et al., 1999, Cho and Harrison, 2012). All other dependencies are therefore indirect and dependent on CBF3s essential role in the localisation of the Cse4 nucleosome specifically to the centromere (explained in detail in 1.4.3). The centromere specific nucleosome is then recognised and bound

by Mif2^{CENP-C} which acts as a central platform to recruit other kinetochore proteins (Carroll et al., 2010, Kato et al., 2013). Besides Mif2, some subunits of the Ctf19^{CCAN} complex also recognise and interact with the centromere-specific nucleosome, a conserved interaction important for Ctf19^{CCAN} stability at the centromere (Ortiz et al., 1999, De Wulf et al., 2003, Carroll et al., 2009). Furthermore, Mif2 directly interacts with components of the Ctf19 and MIND complexes (Cohen et al., 2008, Hornung et al., 2014, Dimitrova et al., 2016). The latter is involved in recruiting the rest of the KMN network, Spc105 and Ndc80 complexes (Dimitrova et al., 2016). This hierarchical assembly of the kinetochore underlies cell-cycle dependent regulation but remains poorly understood (reviewed in McAinsh et al., 2003, Musacchio and Desai, 2017). Whilst the overall architecture and interaction of subunits is conserved, the mechanism by which the centromeric nucleosome is recruited is different depending on centromere type. As budding yeast kinetochore establishment is dependent on CBF3, it is crucial to understand how this complex functions on the centromere.

1.4.1 CBF3 subunits and assembly

The CBF3 complex was discovered in 1991, due to its ability to bind CDEIII in a sequence-specific matter (Lechner, 1991). It consists of four proteins, Ndc10, Cep3, Ctf13 and Skp1 (Figure 8; Lechner, 1991, Goh, 1993, Jiang, 1993, Lechner, 1994, Strunnikov et al., 1995, Doheny et al., 1993, Connelly and Hieter, 1996, Stemmann and Lechner, 1996). All four proteins have been shown to be essential for kinetochore function (Jiang, 1993, Doheny et al., 1993, Strunnikov et al., 1995, Connelly and Hieter, 1996). Both Ndc10 and Cep3 are homodimers in solution, whereas Skp1 and Ctf13, a F-box containing protein, form a heterodimer (Russell, 1999, Purvis and Singleton, 2008, Perriches, 2014). Skp1, the only conserved CBF3 protein, is better known as a subunit of Skp1/Cullin/F-box ubiquitin ligase family (SCF ligases) and also forms part of the RAVE complex (Deshaies, 1999, Seol et al., 2001). Within the SCF ligase Skp1's function is to connect the elongated scaffold subunit Cullin with a variety of F-box proteins (Zheng et al., 2002). The latter bind a range of substrates to be ubiquitinylated and therefore degraded in a proteasome dependent manner (Reitsma et al., 2017). Generally, F-box proteins also contain other protein/protein interaction domains, like WD40 or Leucin-rich repeats (LRRs),

and are classified accordingly (reviewed in Kipreos and Pagano, 2000). Multiple available crystal structures from the budding yeast and human Skp1 bound to different F-box proteins show that the very C-terminus of Skp1 is involved in the binding of the F-box (Figure 8; Schulman et al., 2000, Orlicky et al., 2010, Hao et al., 2005, Kumanomidou et al., 2015, Li and Hao, 2010, Wong et al., 2017). Accordingly, the F-box is necessary for interaction between Ctf13 and Skp1 (Russell, 1999). Besides binding to Skp1 through the F-box domain, Ctf13 also interacts with the other two CBF3 subunits and therefore acts as the central part of CBF3 (Russell, 1999). On its own it is inherently unstable and cannot be recombinantly expressed and purified (Russell, 1999, Perriches, 2014) which accounts for the difficulty in reconstituting CBF3 for functional or structural studies.

Both Ndc10 and Cep3 have been shown to bind to DNA (explained in detail in 1.4.2). Cep3 has a N-terminal Gal4-like C6 zinc cluster (Gal4-domain), which is connected through the rest of the protein through a flexible linker (Lechner, 1994, Strunnikov et al., 1995, Espelin, 1997). A crystal structure, lacking the Gal4-domain, could verify that Cep3 is indeed a homodimer and shares some structural details with HEAT repeats (Figure 8, Purvis and Singleton, 2008, Bellizzi, 2007). Ndc10, the largest subunit of CBF3, contains 956 amino acids. Biochemical evidence showed that Ndc10 forms a homodimer in solution and therefore it compromises half of the mass of CBF3 (Russell, 1999, Perriches, 2014). It is predicted to be highly disordered especially in the C-terminal part, which includes the dimerization and a DNA binding domain, as well as interactions sites for Scm3 and Bir1 (more details below). The crystal structure of the monomeric N-terminal part of the protein (NTD; residues 1 to 550) highlighted a surprising similarity with tyrosine recombinases, although both endonuclease as well as ligase activity is lacking (Figure 8; Perriches and Singleton, 2012, Cho and Harrison, 2012). The co-crystallisation structure of the *K.lactis* Ndc10 NTD equivalent also gave insight into the unspecific DNA binding activity of Ndc10 and suggested that each Ndc10 monomer can bind a separate fragment of DNA, which might facilitate a loop formation at the centromere (Cho and Harrison, 2012). Furthermore, the NTD has been implicated to be important for interaction with other CBF3 subunits (Cho and Harrison, 2012) and the very C-terminus features a degron motif (Furth et al., 2011, Alfassy et al., 2013).

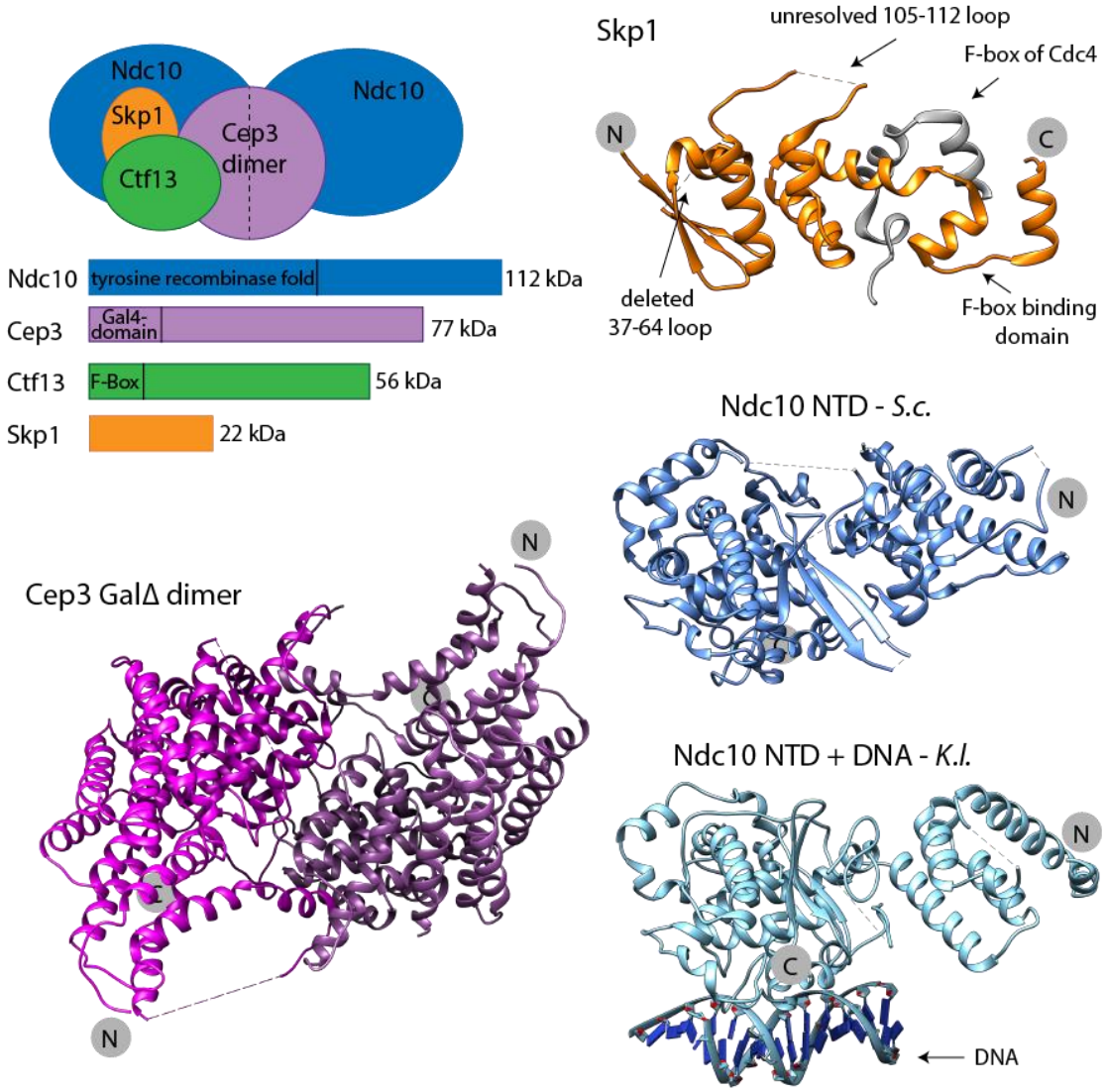

Figure 8: The CBF3 complex and known crystal structures

The schematic model shows the complex in the assumed stoichiometry of 2xNdc10, 2xCep3, 1xCtf13 and 1xSkp1. Known crystal structures include Skp1 (3V7D), Cep3 Gal4Δ (2VEQ) and Ndc10 NTD (*S.c.*: 4ACO; *K.l.*: 3SQI). The construct used for Skp1 had the yeast-specific loop between residues 37 to 64 deleted and was co-crystallised with Cdc4 (F-box depicted in grey). The N-terminal Gal4-domains of Cep3 were deleted to enable crystallisation and the two available crystal structures of Ndc10 NTD are shown: the budding yeast protein without DNA (Ndc10 NTD – *S.c.*) and the *K.lactis* NTD equivalent co-crystallised with DNA (Ndc10 NTD +DNA – *K.l.*).

The formation of the complex is believed to be stepwise occurring and tightly controlled involving phosphorylation, protein degradation and assembly factors (Lechner, 1991, Kaplan, 1997, Russell, 1999). The first step is assumed to be the

formation of the Ctf13/Skp1 heterodimer, which involves the Hsp90-Sgt1 chaperone complex (Kitagawa et al., 1999, Stemmann et al., 2002, Bansal et al., 2004). Sgt1 binds both Hsp90 and Skp1, therefore linking it to both the SCF ligases and CBF3 (Lingelbach and Kaplan, 2004, Willhoft et al., 2017). Interestingly, Sgt1 also specifically targets proteins with LRRs, a common feature of F-box proteins like Ctf13 (Stuttmann et al., 2008, Taipale et al., 2014). Indeed, a weak but direct interaction exists between Sgt1 and Ctf13 (Bansal et al., 2004, Rodrigo-Brenni, 2004). Whilst it is clear that mutations obstructing the Sgt1-Skp1 interaction inhibit CBF3 complex formation, the molecular mechanism remains elusive (Bansal et al., 2004, Kitagawa et al., 1999, Lingelbach and Kaplan, 2004, Willhoft et al., 2017). Whereas most studies seem to conclude that Hsp90 is needed to form the Ctf13/Skp1 heterodimer by somehow activating Skp1 and/or Ctf13, others conclude that the presence of Sgt1 is also needed for further complex formation (Stemmann et al., 2002, Bansal et al., 2004, Rodrigo-Brenni, 2004).

The second step of CBF3 formation is the binding of the Ctf13/Skp1 heterodimer to Cep3 and formation of the core CBF3 complex (Russell, 1999). As unbound Ctf13/Skp1 heterodimer is believed to be readily degraded in a ubiquitin-dependent manner, this step has been referred to as rate-limiting in the formation of CBF3 (Russell, 1999, Kaplan, 1997). Accordingly, Ctf13 has been shown to degrade rapidly *in vivo*, whereas the other subunits are more stable (Kaplan, 1997). Furthermore, an early study suggested that Ctf13 needs to be activated by Skp1-dependent phosphorylation to be able to bind to Cep3 (Kaplan, 1997). This stands in agreement with the discovery that phosphatase treatment of purified CBF3 leads to a loss of protein/DNA complex formation (Lechner, 1991). However, a later study revealed that although Ctf13 and Skp1 contain two and four phosphorylations respectively, complete dephosphorylation of the heterodimer does not interfere with further complex formation (Stemmann et al., 2002). If phosphorylation, possibly of the other two CBF3 subunits, still plays a role in complex formation or activation is unclear.

In the final step of complex formation Ndc10 binds to the core complex forming the DNA-binding active, full complex (Russell, 1999). As discussed in more detail in the next chapter, only this full complex is believed to be centromere-binding sufficient. Besides the full complex, some studies also suggest the existence of an extended complex incorporating a second Ndc10 dimer, which either binds the centromere up- or downstream of the normal CBF3 complex (Espelin, 1997, Espelin, 2003).

1.4.2 Centromere binding of CBF3

CBF3 binds specifically to the CDEIII element of the point centromere (Figure 9). Mutational studies of the centromere revealed that a conserved CCG triplet in this element is essential for centromere function (Ng and Carbon, 1987, Hegemann et al., 1988). Mutations in this CCG triplet strongly affect binding of the CBF3 complex *in vitro* (Ng and Carbon, 1987, Lechner, 1991) and have been shown to interfere with the localization of Ndc10 and Mif2 to centromeric DNA *in vivo* (Meluh and Koshland, 1997).

Early studies have shown that all four subunits of CBF3 are needed to be able to bind the centromere (Sorger et al., 1995, Stemmann and Lechner, 1996, Kaplan, 1997, Espelin, 1997, Pietrasanta, 1999), a surprising result as both Cep3 and Ndc10 can bind DNA on their own (Espelin, 2003, Purvis and Singleton, 2008, Bellizzi, 2007, Perriches and Singleton, 2012, Cho and Harrison, 2012). Furthermore, DNA-protein crosslinking studies showed that only three of the four subunits, namely Ndc10, Cep3 and Ctf13, are in direct contact with centromeric DNA (Figure 9; Espelin, 1997). Whereas there is no evidence that Ctf13 can bind to DNA on its own, Ndc10 has been shown to bind DNA non-specifically (Perriches and Singleton, 2012, Cho and Harrison, 2012). Cep3, on the other hand, accounts for the sequence-specificity involving the conserved and essential CCG triplet of CDEIII and its Gal4-domain (Espelin, 1997, Purvis and Singleton, 2008). Proteins containing this domain, usually bind as a dimer to a tandem or inverted repeat of a CCG triplet with each monomer contacting one triplet (Marmorstein et al., 1992, Marmorstein and Harrison, 1994, Liang et al., 1996, King et al., 1999). Although Cep3 forms a homodimer, only one CCG triplet is present in the centromere. A peculiar characteristic, which has led to a debate of how the homodimeric Cep3 is able to bind this single CCG triplet in the centromere. One hypothesis suggested, that Cep3 contacts a second, degenerate site upstream of the CCG triplet, as has been shown with crosslinking studies (Espelin, 1997). Mutation of this potential TGT triplet, however, did not influence DNA-binding of Cep3 (Purvis and Singleton, 2008) and structural studies will be necessary to answer this question, as well as elucidate how the whole CBF3 complex binds to the centromere.

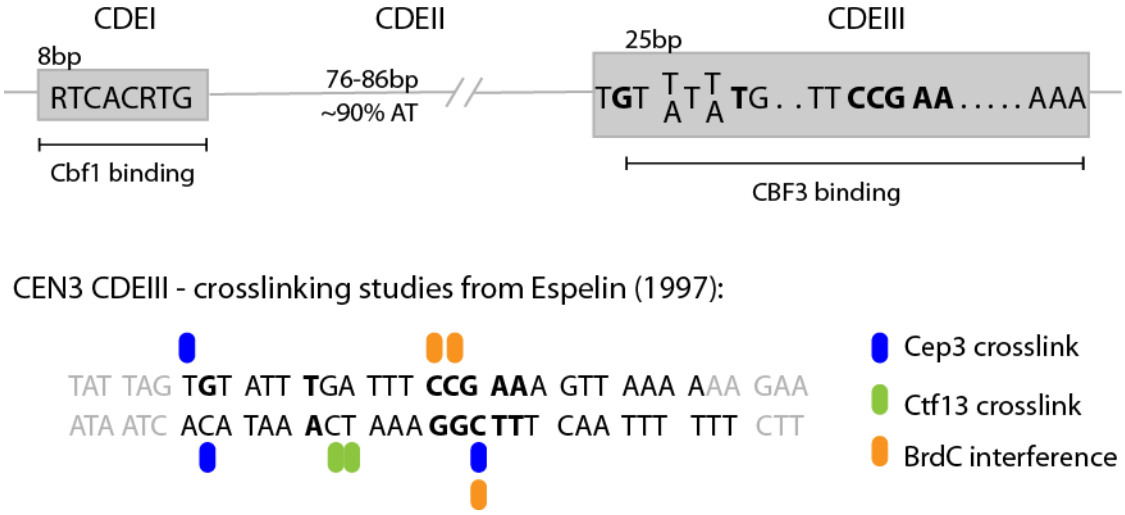

Figure 9: Centromere binding of CBF3

The consensus sequence of the budding yeast point centromere is shown, as well as bases important for CBF3 binding as shown with crosslinking studies (Espelin, 1997).

CBF3 is also believed to have an impact in the overall structure of the point centromere by DNA looping and/or bending. DNA looping, as shown with chromosome conformation assays, is dependent on Ndc10, but not on outer kinetochore proteins (Anderson, 2009, Diaz-Ingelmo et al., 2015). Accordingly, Ndc10 can bind two individual stretches of DNA *in vitro* and an atomic force microscopy study has suggested that CBF3, upon binding, bends the centromeric DNA by about 55°, as well as being able to bind multiple DNA fragments (Cho and Harrison, 2012, Pietrasanta, 1999). Although CBF3 proteins are not conserved outside point centromere species, this DNA bending could depict a conserved function to the mammalian inner centromere protein CENP-B. The crystal structure of CENP-B DNA binding domain in complex with DNA has shown, that it too introduces a bend in the DNA, which could be important for its CENP-A nucleosome assembling ability (Okada et al., 2007).

Another interesting aspect of CBF3 DNA binding is the fact that it needs to accommodate the additional proteins also found to bind to the short centromeric DNA (see 1.3.1). How all of these proteins simultaneously bind and possibly interact is still unclear and further studies are needed.

1.4.3 Interaction with Scm3 and Cse4 loading

The localisation of Cse4 to the centromere is dependent on CBF3, although there is no direct interaction between the two (Measday et al., 2002, Ortiz et al., 1999). The domain essential for centromere targeting of $Cse4^{CENP\text{-}A}$ is the conserved CATD domain within the histone-fold (Black et al., 2004, Wieland et al., 2004). This domain is sufficient in human, whilst in budding yeast a short stretch of the long N-terminal tail of Cse4 is also needed for centromere targeting (Chen et al., 2000, Keith et al., 1999). Both, however, are specifically bound by the conserved, centromere-specific histone chaperone $HJURP^{Scm3}$, an interaction essential for centromere localisation (Camahort et al., 2007, Stoler et al., 2007, Aravind et al., 2007, Sanchez-Pulido et al., 2009, Shuaib et al., 2010). Biochemical and structural studies showed a conserved interaction between the N-terminal part of $HJURP^{Scm3}$ and the CATD of $CENP\text{-}A^{Cse4}$ (Foltz et al., 2009, Hu et al., 2011, Zhou et al., 2011, Cho and Harrison, 2011, Figure 10).

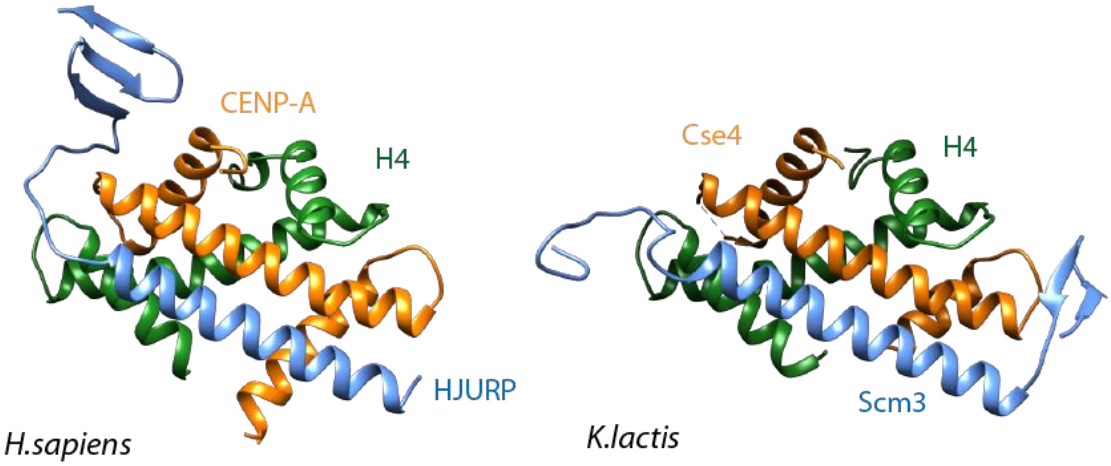

Figure 10: Conserved Scm3/Cse4/H4 interaction

Crystal structures of the human HJURP/CENP-A/H4 interaction (left; 3R45) and the *K.lactis* homologues, Scm3/Cse4/H4 (right; 2YFV).

In regional centromere species an elaborate process of centromere licensing, which involves specific histone modifications, the presence of CENP-A, the function of the Mis18 protein complex as well as additional histone chaperones, is responsible for proper incorporation of new CENP-A (reviewed in Stellfox et al., 2013, Musacchio and Desai, 2017). In point centromere species, on the other hand, the conserved centromeric sequence and CBF3 are sufficient for correct Cse4 localisation (Ortiz et

al., 1999, Measday et al., 2002). This simpler process is believed to be facilitated through a direct interaction between the CBF3 subunit Ndc10 and the Cse4-specific chaperone Scm3 (Camahort et al., 2007, Cho and Harrison, 2011). However, the molecular mechanism of this interaction, as well as how it facilitates the incorporation of Cse4/H4 into the centromeric nucleosome and in fact the exact composition of this nucleosome, discussed in 1.2.1, remains unclear.

1.4.4 Additional functions

Beside its main function in specifying the centromere for kinetochore establishment and building a platform for kinetochore formation, CBF3 also is involved in the regulation of kinetochore function, as well as spindle stability.

Involvement in SAC signalling: Most *S.cerevisiae* kinetochore mutants, which are defective in kinetochore-microtubule attachment, trigger the SAC and therefore lead to cell cycle arrest (Gardner, 2001). Ndc10-1 mutants, however, still complete anaphase with massive chromosome mis-segregation, although the cell fails to establish a functional kinetochore (Gardner, 2001, Goh, 1993). This, and the fact that they cannot maintain an arrest in the presence of microtubule-polymerization inhibiting drugs (Tavormina and Burke, 1998), shows that they are not checkpoint proficient. Accordingly, the CBF3 subunit Skp1 has been shown to directly interact with the SAC protein Bub1. This interaction, shown to be independent of Skp1's function in the SCF ligases, recruits Bub1 to the centromere (Kitagawa et al., 2003). Furthermore, a specific Skp1 mutant could be constructed, which is deficient in Bub1 interaction but does not affect chromosome segregation. This mutant was hypersensitive to microtubule destabilising drugs, a typical response of SAC checkpoint deficiency (Kitagawa et al., 2003). These studies suggest that CBF3 is important to recruit SAC proteins through its subunit Skp1.

Involvement into CPC localisation: CBF3, through its subunit Ndc10, also interacts with Bir1, a component of the CPC (Cho and Harrison, 2012, Yoon and Carbon, 1999). Bir1 therefore is coupled to the centromere, which has been shown to be crucial for CPC localisation and function in error correction (Shimogawa et al., 2009, Makrantoni and Stark, 2009). Interestingly, a different mechanism exists in regional centromere species, which lack CBF3, involving Shugoshin, which recruits Bir1 and specifically recognises phosphorylated H2A (Kawashima et al., 2010).

Involvement in spindle stability and cytokinesis: The first evidence of a CBF3 function outside of the centromere, came from the finding that Ndc10 moves to the spindle midzone during late anaphase and is important for spindle stability and cytokinesis (Bouck and Bloom, 2005). Interestingly, the same is true for the CPC (Thomas and Kaplan, 2007). As explained above, Ndc10 directly interacts with CPC through Bir1, which together with the identical localisation pattern indicates a co-migration of both complexes to the spindle midzone (see above; Thomas and Kaplan, 2007, Cho and Harrison, 2012, Gillis et al., 2005). Furthermore, defects in assembly or turnover of the CBF3 complex have been shown to cause defects in cytokinesis, indicating that not only Ndc10, but the whole CBF3 complex is involved (Gillis et al., 2005). Again, Bir1 mutant studies revealed similar defects, arguing for a co-functionality as well as co-migration of these two complexes (Gillis et al., 2005). It is unclear how CBF3 localisation to the spindle midzone is regulated.

Involvement in DNA-damage signalling: After sensing of DNA double-strand breaks the DNA-damage response increases chromatin mobility, a process believed to promote repair by homologous recombination (Krawczyk et al., 2012, Lottersberger et al., 2015, Dion et al., 2012, Mine-Hattab and Rothstein, 2012). In yeast the kinase Mec1 is essential for this process and one downstream target of Mec1 activity is the CBF3 subunit Cep3 (Dion et al., 2012, Mine-Hattab and Rothstein, 2012, Strecker et al., 2016). Cep3 is specifically phosphorylated on the S575 residue by Rad53, which is activated in a Mec1 dependent manner upon DNA damage. This phosphorylation seems to lead to detachment of kinetochore/spindle microtubule, which in turn increases chromatin mobility (Strecker et al., 2016).

1.5 Project aims

The aim of this thesis was the functional and structural characterisation of the budding yeast CBF3 complex. Being the initial factor of budding yeast kinetochore establishment, it is crucial to understand how this complex functions on the centromere and therefore influences budding yeast kinetochore establishment. For reasons described above budding yeast is one of the main model systems used for kinetochore studies. Structural and functional studies of CBF3, as well as attempts to reconstitute the full kinetochore have suffered from difficulties in recombinant

expression and purification of CBF3. An initial aim therefore was to overcome these difficulties and find a working recombinant expression and purification system for the full CBF3 complex. This was followed by structural, as well as functional studies to elucidate its composition, centromere binding activity, as well as interactions with other kinetochore proteins, especially its interaction with Scm3 and influence on Cse4 loading.

1.6 Theory of single particle cryoEM

To fully understand how a protein or protein complex folds, interacts and therefore functions structural biology is key. In the past X-ray crystallography was the workhorse for structural biology providing 3D-structures helping to understand how proteins work within the cell. The biggest limitation of this method, however, is the fact that one must produce a protein crystal to be able to achieve this high-resolution information and for many proteins and especially macromolecular assemblies this can be tricky or impossible (Egelman, 2016). Furthermore, a protein is stabilised in one particular conformation in the crystal lattice, providing therefore only one snapshot of a protein's life. In recent years a new technique, cryo electron microscopy (cryoEM), is gaining more and more importance in structural biology, recognised last year by the award of the Nobel Prize in Chemistry to Jacques Dubochet, Joachim Frank and Richard Henderson (Nogales, 2018). CryoEM has a few significant advantages over crystallography. First, no crystallisation is needed, which opens up structural studies of proteins and big macromolecular assemblies, which have proven to be difficult to crystallise, as well as those which can only be purified in low yields. Second, proteins are imaged in solution, which does not restrict their conformational states and therefore multiple structures can be determined from one sample describing the function of the protein better than only one snapshot (Cheng et al., 2015, Scheres, 2016). Third, it is possible to image bigger macromolecular assemblies directly in the cell and still achieve low to medium resolution structures, which never would have been possible with crystallography (Jonic et al., 2008, Briggs, 2013, Oikonomou et al., 2016).

Many achievements in recent years, including progress in detector quality, motion correction and image processing software, have given rise to a 'resolution-revolution'

with many cryoEM structures reaching a resolution below 3.5Å (Henderson and McMullan, 2013, Kuhlbrandt, 2014, Subramaniam et al., 2016, Vonck and Mills, 2017, Liao et al., 2013). Furthermore, with resolutions good enough to visualise small molecules, like water or inhibitors, cryoEM is gaining momentum in the pharmaceutical industry for structure-based drug design (Boland et al., 2017, Zhu et al., 2018, Khoshouei et al., 2017, Fischer et al., 2015, Bartesaghi et al., 2018).

1.6.1 The transmission electron microscope (TEM) – image formation

The idea of a transmission microscope using electrons instead of light to form a higher resolution image comes from the Abbe's equation. This equation specifies that the highest resolution obtainable ultimately is dependent on the wavelength of the imaging radiation. Therefore, it is not possibly to resolve anything under ~200nm with a light microscope. Many biological samples including proteins and protein complexes, however, are much smaller than this. Electrons, on the other hand, depending on their energy exhibit a wavelength in the picometer scale, enabling modern TEMs to reach a resolution of better than 1Å (Smith, 2008).

For image formation electrons need to be accelerated to form a parallel, high energy beam, which then can interact with the sample leading to scattering of the said electrons. Two different forms of scattering can occur: elastic and inelastic scattering. Elastically scattered electrons only change their moving direction but not their energy as they interact with an atom of the specimen. Depending on the scattering angle, they either are excluded by the objective aperture and therefore contribute to amplitude contrast (high angles) or they are focused back to their original point by the objective lens, where they interact either positively or negatively with the unscattered beam and therefore produce phase contrast (low angles). Whilst perfectly focussed images produce no phase contrast, the higher the defocus the better the phase contrast of low resolution details of the image. However, high defocus also leads to a loss of high resolution details, giving rise of the need of careful balancing of contrast and high-resolution information loss. Inelastic scattering occurs when the incoming electron hits an electron from the specimen, which leads not only to a change of moving direction but also to a change in energy. These electrons are removed by an energy filter and therefore contribute to amplitude contrast. As electrons interact very weakly with a biological sample, the overall

amplitude contrast is negligible and phase contrast, produced by defocusing, is the main contributor to the image (Orlova and Saibil, 2011, Frank, 2006).
Magnification of the electrons forming the image is achieved by electromagnetic lenses. These lenses consist of a wire coil through which current flows. The current creates a magnetic field, which influences the travel path of electrons opening the possibility to focus and magnify in the same way an optical lens does. A TEM usually has multiple electromagnetic lenses to magnify the image, which together with deflectors form the projector lens system. Through adjusting the current which flows through each of these lenses, one can adjust the magnification. Besides the projector lenses a TEM also incorporates an objective lens system and a condenser lens system, to facilitate focussing of the image, as well as condensing the beam onto a specific area of the specimen respectively (Frank, 2006, Orlova and Saibil, 2011).

1.6.2 Sample preparation

One of the difficulties to overcome in imaging biological samples with a TEM is that electrons can only travel in high vacuum. Therefore, the sample needs to be prepared in a way to withstand these conditions. Traditionally protein or other biological samples, like viruses, could only be imaged by staining and subsequent drying of the sample. As outlined below, this limits the achievable resolution to about 20Å (Ohi et al., 2004). Another possibility to conserve the sample for use in the high vacuum is freezing it and imaging at low temperature to maintain the frozen state. This, however, was complicated by the fact that frozen water molecules are crystalline, which both damages biological molecules and produces a high background in the TEM image. An important achievement, proposed in 1940 and finally realised in 1980/81, was the vitrification of water (Brüggeller and Mayer, 1980, Dubochet and Booy, 1981, Dubochet and McDowall, 1981). The idea is simple, cool water rapidly enough that the molecules have no time to crystallise. Vitrified water, in contrast to the crystalline form, is amorphous and this glass-like appearance is key for both maintaining a biological molecule in its native state as well as being able to visualise it without overwhelming background. Up to now, both these methods of sample preparation are crucial for structural studies with cryoEM and are explained in more detail below.

Negative stain: Proteins or protein complexes, given they fulfil a minimum size requirement, can be quickly and easily visualised in a TEM by using a heavy metal stain. Carbon coated grids are used as a support for the protein sample to adhere to. The stain fixes the protein and after blotting away the stain and drying the grid, a thin layer stays behind outlining the protein particles (De Carlo and Harris, 2011). The micrograph obtained from such a grid, will show the stain as dark background and particles appear lighter surrounded by a dark contour. As one is imaging the stain rather than the protein particle, only the overall shape of protein can be visualised and no high-resolution information can be obtained. Also, the protein will be flattened to some degree as it loses its hydration shell when the grid dries (Kiselev et al., 1990). However, negative stain is very useful, if not essential, in judging particle quality and concentration and can be used to obtain initial low-resolution models (Passmore and Russo, 2016).
Grids for negative stain imaging can be coated with carbon in-house or can be bought already coated with carbon. Before use grids are glow discharged or plasma cleaned. This changes the surface of the grids from hydrophobic to hydrophilic, and therefore allows protein solution to be pipetted on the grid and protein molecules to adhere on the carbon surface (Aebi and Pollard, 1987).

Cryo freezing: For high resolution structure determination, a protein or protein complex needs to be vitrified in a thin layer of buffer to retain its natural hydrated state and to be able to image it in the high vacuum of the TEM. To achieve this, the protein solution is added on a carbon or gold grid with a holey support film and after excess is blotted away it is rapidly frozen by plunging it into liquid ethane cooled to liquid nitrogen temperatures (Dubochet et al., 1988, Passmore and Russo, 2016). On an ideal grid, the protein is frozen in random orientations in a thin layer of ice within the holes of the grid, whereas the ice thickness is only slightly thicker than the protein specimen itself. Thin ice is preferable as it ensures a good signal-to-noise ratio (Orlova and Saibil, 2011). Many specimens, however, exhibit unwanted behaviour on the air-water interphase, such as preferred orientation or complete degradation (Glaeser and Han, 2017). To avoid these problems, one can try different adjustments in sample preparation including imaging in thicker ice, using a carbon or graphene support or including surfactants in the buffer to protect the specimen from the air-water interface. Once the best freezing conditions are established it is

advisable to collect a small dataset on a low-end microscope to establish if the specimen shows preferred orientation, before collecting on a high-end microscope.

1.6.3 Direct electron detectors and data collection

A major contributor to the 'resolution-revolution' was the advent of direct electron detectors in 2012 (Kuhlbrandt, 2014, McMullan et al., 2014, Bai et al., 2015). Before, either film or scintillator-based digital cameras such as charge-coupled device (CCD) were used for data collection. Film was generally preferred, as it performs better in recording high-resolution frequencies than a CCD camera, measured as detective-quantum-efficiency (DQE) and modulation transfer function (MTF). Image acquisition, however, is slow and high defocus is needed to achieve enough contrast (Wu et al., 2016). Direct detectors have two main advantages. First, their DQE and MTF is generally significantly higher than either film or CCD cameras and second, due to the high frame rate of direct detectors it is possible to record multiple frames per exposure, also called a movie stack (Campbell et al., 2012, Li et al., 2013, McMullan et al., 2014, Ruskin et al., 2013, Zheng et al., 2017, Scheres, 2014). The latter is especially important as it allows to correct for beam-induced movement of the specimen, as well as weighting of high- and low-resolution frequencies of different frames according to radiation damage (Brilot et al., 2012, Bai et al., 2015). Radiation damage is one of the limitations of cryoEM, as the imagining electrons severely destroy biological samples in a short frame of time (Baker and Rubinstein, 2010).

CryoEM data for single particle analysis (SPA) is collected without tilting the specimen, a method used for tomography (Jonic et al., 2008). An automated software is used to collect data for multiple days to reach the required number of particles (Tan et al., 2016). Typically, a pixel size of ~1Å is chosen, which theoretically allows to resolve features as small as ~2Å, which represents the Nyquist frequency. Dose rate should be optimised depending on camera used and parameters such as defocus value, use of a phase plate, camera type are dependent on the specimen, especially its size (Wu et al., 2016).

1.6.4 Data processing

High-resolution structure determination with cryoEM would not be possible without sophisticated image processing software. The first step in a typical SPA workflow has already been mentioned and is the global motion correction of the movie stack, as well as batch correction for local movements and weighting according to radiation-damage. To minimise processing time, one can easily do these corrections on-the-fly as soon as movies are collected using automated software like for example Scipion (de la Rosa-Trevin et al., 2016).

Subsequently, another correction according to the contrast transfer function (CTF) has to be done (Penczek et al., 1997, Zhang, 2016, Mindell and Grigorieff, 2003, Rohou and Grigorieff, 2015). This originates from the way a TEM modifies and therefore distorts an image (Frank, 2006). Depending on the exact defocus value it over- or underrepresents values as a function of spatial frequency, losing some information altogether. By estimating the exact defocus value, software can correct for this effect and by collecting images at different defocus values lost information can be restored.

Once images are corrected, particles need to be recognised and picked, a process carried out automatically due to the typical large size of datasets (up to millions of particles). Sophisticated SPA reconstruction algorithms align and compare these randomly orientated particles and group them depending on their similarity, a process called 2D classification. As particles of one group are averaged together the signal-to-noise ratio is significantly improved. These 2D projections are then used to iteratively reconstruct the original 3D shape, by determining the unknown orientations of the 2D projections (Frank, 2006, Orlova and Saibil, 2011). Classification, both in 2D and 3D, is also a powerful tool in sorting out junk particles and in differentiating between different conformations of heterogeneous datasets. The most used software packages implementing all steps from particle picking to 3D reconstructions are Eman, Relion and CryoSparc (Tang et al., 2007, Scheres, 2012, Punjani et al., 2017).

Chapter 2. Materials & Methods

2.1 Molecular Cloning

2.1.1 Restriction-free cloning

Restriction-free (RF) cloning is a simple method to insert a DNA fragment into any position of a DNA plasmid without use of restriction enzymes and ligation (van den Ent and Lowe, 2006). The sequence one wishes to insert, is amplified with overhangs, which are complementary to 5' and 3' of the insertion site of the plasmid. This so-called megaprimer is then used in a second PCR, which amplifies around the whole plasmid. Primers can be automatically designed using a webserver (Bond and Naus, 2012). One can use genomic, cDNA or another plasmid containing the gene of interest as a template and start/stop codon can be removed or added by simply changing the primer for the first PCR. The megaprimer is gel purified (QUIquick Gel Extraction Kit, Qiagen) before being used in the second reaction. Dpn1 treatment is used to digest parental plasmid DNA for 1hr at 37°C, after which 5μl is transformed into chemical competent XL1 cells. Cells are plated on the appropriate antibiotic-containing plates and grown overnight. Single colonies are picked and insertion is tested by colony PCR (see 2.1.6). ONCs are grown from positive clones and plasmid DNA is extracted (Quigen Miniprep kit) and sequenced (see 2.1.7).

1st PCR	
5x Phusion HF Buffer	10μl
dNTPs (10mM)	5μl
fw primer (10μM)	2μl
rv primer (10μM)	2μl
template (30-50ng/μl)	1μl
Phusion polymerase	1μl
H_2O	29μl
Total volume	50μl

2nd PCR	
5x Phusion HF Buffer	10μl
dNTPs (10mM)	2.5μl
plasmid (40ng/μl)	1μl
megaprimer*	x μl*
Phusion polymerase	1μl
H_2O	x μl**
Total volume	50μl

* Molar ratio of plasmid to megaprimer 1:40

**up to 50μl

PCR program		
98°C	30sec	
98°C	30sec	
55°C	30sec	30x
72°C	~30sec/kbp	
72°C	~1min/kbp	
4°C	∞	

PCR program		
98°C	30sec	
98°C	30sec	
50°C	1min	18x
72°C	12min*	
4°C	∞	

* unless over 6kbp, then 2min/kbp

2.1.2 Gibson assembly cloning

Gibson assembly cloning (Gibson et al., 2009) can be used to join two or more DNA fragments together, independent of restriction enzyme sites or sequence. It uses overlapping DNA fragments, produced by PCR, and a cocktail of three different enzymes: an exonuclease, to create single strand 3' overhangs; a polymerase to fill in the gaps in the annealed DNA; and a DNA ligase which joins annealed fragments covalently.

Both commercially available Gibson Assembly® (NEB) and NEBuilder® HiFi DNA Assembly Cloning kit (NEB) were used as specified in their protocols. Plasmid was either linearized with restriction enzymes (if suitable ones were present) or via PCR, and primers were designed as described in the protocols.

Insert/Vector-linearization PCR	
5x Phusion HF buffer	10µl
dNTPs (10mM)	5µl
fw primer (10µM)	2µl
rv primer (10µM)	2µl
template (50ng/µl)	1µl
Phusion polymerase	1µl
H_2O	29µl
Total volume	50µl

PCR program		
98°C	30sec	
98°C	30sec	
55°C	30sec	25x
72°C	~30sec/kbp	
72°C	~1min/kbp	
4°C	∞	

2.1.3 Traditional cloning

In traditional cloning a gene of interest can be inserted into a multiple cloning site (MCS) of a plasmid, by using restriction enzymes to create small complementary overhangs, which then can be annealed and ligated. The insert therefore has to be amplified with primers adding the desired restriction site, to enable to cut both plasmid and insert with the same restriction enzymes. Despite being a reliable method, it can only insert DNA fragments at existing restriction sites in the plasmid, but nowhere else. Also, a seamless insertion of a DNA fragment of interest cannot be achieved. This method was used, however, when cloning genes into the bidirectional yeast expression vectors, as there was the need to insert genes between flanking sequences identical at both MCSs. This made above mentioned cloning methods unusable.

Insert PCR		**PCR program**		
5x Hi-Fi Buffer	10µl	98°C	2min	
dNTPs (10mM)	5µl	98°C	30sec	
fw primer (10µM)*	2µl	53°C	30sec	25x
rv primer (10µM)*	2µl	72°C	~30sec/kbp	
template (30ng/µl)	1µl	72°C	~1min/kbp	
Velocity polymerase	1µl	4°C	∞	
H_2O	29µl			
Total volume	50µ			

* Primers contain required restriction sites and overhangs

The quality and quantity of PCR product is checked by agarose gel electrophoresis, and subsequently purified (QUIquick PCR purification kit, if amplification yielded single band, or QUIquick gel extraction kit, if contamination bands are visible). Both insert and plasmid is digested with appropriate restriction (NEB) enzymes for 2-4hrs at the applicable temperature. Cut plasmid is then treated with alkaline phosphatase (CIP, NEB) to avoid re-ligation, and both insert and plasmid is separated via agarose gel electrophoresis and purified using the QUIquick gel extraction kit. T4 DNA ligase is used to ligate insert and plasmid at RT for at least 4hrs or ON, and ligation mix is then transformed into XL1 cells, checked by colony PCR and sequenced.

Restriction digest	
Enzyme A	2µl
Enzyme B	2µl
Insert or plasmid	~5µg
Appropriate buffer*	3µl
H_2O	xµl**
Total volume	30µl

Ligation reaction	
Plasmid	xµl***
Insert	xµl***
T4 Ligase buffer	0.5µl
T4 DNA Ligase	0.5µl
H_2O	xµl**
Total volume	5µl

* a buffer is chosen in which both enzymes exhibit 100% activity. If a double digest is not possible, insert and plasmid are cut subsequently.

** appropriate amount to make up to total volume

*** Plasmid and insert are used in a molar ratio of 1:3

2.1.4 Site-directed mutagenesis

Site-directed mutagenesis creates specific point mutations in DNA, to establish single or double mutations in protein sequence or mutations in DNA sequence (e.g. centromere, promoter, etc.) The method, described by Carey et al., 2013, involves a PCR reaction with complementary primers, which contain the desired mutations. Primers were designed using the webtool PrimerX (http://www.bioinformatics.org/primerx/). Parental DNA was digested by Dpn1 at 37°C for 1hr after the PCR reaction, and the product is then separated with agarose gel electrophoresis to confirm the size. Subsequently, the mix is directly transformed into XL1 cells, plated on selective plates, and mutations are verified by sequencing.

Mutagenic PCR	
10x Pfu Buffer	5µl
dNTPs (10mM)	2.5µl
fw primer (10µM)	2µl
rv primer (10µM)	2µl
template (50ng/µl)	0.5µl
Pfu polymerase	1µl
H_2O	37µl
Total volume	50µl

PCR program		
95°C	1min	
95°C	30sec	
55°C	1min	18x
68°C	~2min/kbp	
72°C	~1min/kbp	
4°C	∞	

2.1.5 Deletion PCR

A PCR-based deletion method described in Hansson et al., 2008, was used to create constructs with deletions of any size. The method uses only one round of PCR with primers designed exactly starting and ending at the last/first base outside the DNA to be deleted. Therefore, it is a sequence independent and simple method. To facilitate re-ligation of the PCR product, a blunt-end creating polymerase is used and parental plasmid DNA is digested using Dpn1 for 1hr at 37°C. DNA is then purified using the QUIquick PCR purification kit (Qiagen) and DNA product length is verified with agarose gel electrophoresis. 5' ends are phosphorylated by treatment with T4 polynucleotide kinase for 1hr at 37°C. T4 DNA ligase is added to the mix and incubated for at least 3hrs, or overnight at RT. Ligated DNA is subsequently transformed in XL1 cells, plated on selective plates and single colonies are tested by colony PCR and sequencing (see 2.1.6 and 2.1.7).

Deletion PCR	
5x Hi-Fi Buffer	10µl
dNTPs (10mM)	5µl
fw primer (10µM)	2µl
rv primer (10µM)	2µl
template (30ng/µl)	1µl
Velocity polymerase	1µl
H_2O	29µl
Total volume	50µl

PCR program		
98°C	1min	
98°C	30sec	
55°C	30sec	22x
72°C	~1min/kbp	
72°C	~2min/kbp	
4°C	∞	

5' Phosphorylation /Ligation reaction	
PCR product (purified)	8µl
T4 ligase buffer (with ATP)	1µl
T4 Polynucleotide kinase	1µl
Total volume	10µl

→ incubate 37°C for 1 hr; then add 1µl T4 DNA ligase and incubate 3hrs-ON at RT

2.1.6 Colony PCR

Colony PCR is a quick method to screen for successful insertion of the desired DNA piece after cloning. There is no need for plasmid purification, as cells can be directly added to the PCR mix to amplify the region which should contain the insert. ONCs can then be grown only from positive clones, for subsequent plasmid purification and sequencing. Colony PCR can be performed on single clones, or on multiple clones. The latter being useful if cloning efficiency is low. MangoTaq DNA polymerase (Bioline) was used, as it allows to directly load the PCR mix onto an agarose gel after amplification to further speed up the whole procedure.

Colony PCR	
10x MangoTaq buffer	5µl
$MgCl_2$ (10mM)	1µl
dNTPs (10mM)	2µl
fw primer (10µM)	1µl
rv primer (10µM)	1µl
MangoTaq polymerase	0.75µl
H_2O	14.25µl
Total volume	25µl

→ add a small amount of cells with a sterile pipette tip or inoculating loop

PCR program		
94°C	5min	
94°C	30sec	
53°C	30sec	22x
72°C	~30sec/kbp	
72°C	~1min/kbp	
4°C	∞	

2.1.7 DNA Sequencing

DNA sequencing was carried out by the Francis Crick Institute's Genomics Equipment Park, using the BigDye™ Terminator Cycle Sequencing kit (Thermo Fisher Scientific) and capillary Sanger sequencing on an ABI Prism 3730. Sequencing results were compared against the sequence of interest with SnapGene software to exclude any mutations as well as check on correct in-frame insertion.

2.1.8 Agarose gel electrophoresis

Agarose gel electrophoresis is a method used to separate DNA fragments according to their size by applying an electric field. The negatively charged DNA migrates through an agarose gel towards the positive pole. Small fragments pass through the gel faster than big ones. Ethidium bromide is used to detect DNA, as it intercalates into the major groove and fluoresces when excited with UV light. After separation, it is possible to re-extract and purify DNA bands from the gel (QUIquick gel extraction kit, Qiagen). Depending on the sizes of DNA one wishes to separate, the concentration of agarose lies between 0.8%-2%. Agarose is dissolved in an appropriate volume of Tris/Borate/EDTA (TBE) buffer by heating in a microwave. The mixture is left to cool, before adding ethidium bromide solution (10mg/ml, 1:25k dilution) and a gel is cast in the running chamber. After the gel has solidified, DNA samples are prepared with 5x DNA loading buffer blue (Bioline) and run with 80V for 20-30min. DNA is visualised with UV light in a Gel Doc™ EZ Imager (BioRad).

2.2 Bacterial protein expression

Bacteria are the simplest of the available expression systems. Time needed from cloning to testing of constructs is short and growing bacteria is also cheaper than other expression systems. Cells can be grown to high density and mutant, expression-optimised strains are commercially available. However, not all higher eukaryotes' proteins can be successfully expressed, as bacteria cells cannot perform all posttranslational modifications (PTMs) and they often lack cofactors needed to properly fold the protein of interest.

2.2.1 Transformation

Expression plasmid with desired gene(s) under the control of a T7 promoter were transformed into chemically competent *E.coli* BL21 (DE3) cells, by adding 0.5μl of ~100ng/μl plasmid to 50μl of cells. The mixture is incubated for 20min on ice, before being heat shocked at 42°C for 45sec. The cells are briefly put back on ice and 700μl SOC media is added. Cells are then grown for 1hr at 37°C, 300rpm and subsequently plated on LB-Agar plates containing the appropriate antibiotic. Co-expression of two

genes was done using an expression plasmid with two MCSs, both under the control of a T7 promoter. Co-expression of more than three genes was achieved by co-transformation of multiple plasmids, carrying different selective markers and plating on LB-Agar plates containing all these antibiotics.

2.2.2 Expression

An overnight culture (ONC) is grown from one single clone of a fresh transformation in LB-media with the appropriate antibiotic(s). Expression in bacteria can be done at different temperatures; 37°C, the optimal temperature for growth of *E.coli*, or at lower temperature, usually between 16°C to 18°C. The latter can help to properly fold the protein of interest, as it provokes a slowed down metabolism. However, slightly toxic proteins, such as histones, need to be expressed quickly at 37°C.

Expression at 37°C: 200-600µl of ONC is added to one litre of the expression culture, grown in LB-media with added antibiotics. A total of 2-10 litre, depending on expression levels, are used per expression. The expression culture is grown at 37°C and 220rpm to an OD_{600} of 0.6 and subsequently it is induced by adding 0.5ml of 0.5M IPTG to every litre of culture. Protein is expressed for 4hrs, at 37°C and 220rpm, before being harvested via centrifugation at 3000g, 4°C, 10min.

Expression at 16°C: 100µl of ONC is added per one litre of expression culture at noon and the cells are grown to an OD_{600} of 0.4 at 37°C, 220rpm. Temperature is then lowered to 16°C, and after one-hour cells are induced by adding IPTG, as described above. Protein is expressed for 16hrs, before harvesting the cells through centrifugation, as mentioned above.
To wash the cells from residual expression media, they are resuspended in a small amount of appropriate lysis buffer and pelleted again through centrifugation. Cell pellets are then either flash-frozen in LN_2 or processed immediately.

2.2.3 Cell lysis

Cells are resuspended in 5-10 times the amount of lysis buffer (e.g. 5g of cells to 25-50ml of lysis buffer), supplemented with EDTA-free protease inhibitor cocktail set III (Millipore) and lysed through sonification. Therefore, cell suspension is placed on wet ice, to allow for sufficient cooling during the 6x30sec of sonication at step 6 with

a Branson Sonifier 450. Between each pulse cells are rested for 30sec, to avoid overheating. The lysate is subsequently centrifuged at 30,000g for 50min, to separate soluble from insoluble proteins and from cell debris. Supernatant was transferred into a fresh beaker and filtered with a 5μm syringe filter.

2.2.4 Protein expression trials

Due to high expression yields in bacteria, one can test the expression levels in a small-scale culture, without any purification necessary. The lysate can simple be separated on an SDS-PAGE, to check if the recombinant protein is expressed and compare level of expression between different induction temperatures or other factors (e.g. length of induction, IPTG concentration, cell density or between different constructs). To do this, an ONC is grown from a single colony in 5ml of LB-media supplemented with the appropriate antibiotic(s). A 10ml expression culture is inoculated to an OD_{600} of 0.3 with a suitable amount of ONC. For an induction at 37°C, the culture is grown until it has reached an OD_{600} of 0.6 and is then induced with a final concentration of 250μM IPTG and protein is expressed for 4hrs, before cells are harvested via centrifugation and lysed via sonification. For an induction at 16°C, cells are grown to an OD_{600} of 0.5, then cooled to 16°C and induced once they reached an OD_{600} of 0.6, as described above.

2.3 Budding yeast protein expression

Many yeasts can be used as an expression system for recombinant eukaryotic proteins. It is easy to handle and quick to genetically modify. It can perform posttranslational modifications, such as phosphorylations, which are frequent on eukaryotic proteins. Yet, it is relatively cheap to grow and expression can easily be upscaled by using fermenters. Whilst several yeast expression systems exist from various genera, like *Pichia*, *Kluyveromyces* or *Saccharomyces*, budding yeast itself was chosen here to express the CBF3 proteins and to ensure proper complex formation.

2.3.1 Lithium acetate yeast transformation

Generally, yeast expression plasmids can either be self-propagating or integrated into the genome. The latter was used in these studies and they were kindly provided from John Diffley. Each of these plasmids has a bidirectional Gal1-10 promoter to accommodate two genes of interest and a selective marker. The expression strain, also provided from John Diffley, has several amino acid auxotrophies. Once transformed, the selective marker can complement for one auxotrophy, meaning positive clones can survive on minimal media without addition of one particular amino acid, whereas the original strain cannot. To facilitate integration, the plasmid is cut somewhere in the selective marker and introduced into the cells, which will integrate the plasmid spontaneously via homologous recombination. The lithium acetate (LiAc) method was used to introduce the plasmid DNA into the cells. Lithium acetate, in combination with PEG makes the cell wall porous to facilitate DNA uptake, which requires a single-strand carrier DNA, as well as the plasmid DNA of interest. To make the competent cells, a 5ml ONC is grown in YP-media with added 2% Glucose as carbon source. A 50ml day culture (YP-media plus glucose) is inoculated with 600µl of ONC and grown for 4hrs. Log phase cells are then harvested via centrifugation at 2000g, washed once with water and once with freshly made LiAc/TE buffer. Cells are pelleted and resuspended in a final amount of 150-200µl of LiAc/TE buffer. 2µg of cut plasmid (the 20µl restriction digest reaction can be used directly) is mixed with 5µl of salmon sperm DNA, which has been boiled for 5min and then cooled down on ice. 50µl of cells are added to the DNA mix, vortexed for 5sec, before adding 300µl of LiAc/TE/PEG solution and vortex for another 5sec. The mix is then incubated for 30min at 30°C 300rpm and subsequently heat-shocked at 45°C for 15min. Cells are briefly put on ice, before being gently pelleted at 1500g for 30sec. The supernatant is removed and cells are gently resuspended in 300µl of 1M Sorbitol. 100µl of cell suspension is plated per YN-Agar plate lacking the appropriate amino acid. As a control, 50µl of cells are transformed with carrier DNA only, and plated on two plates, one lacking the amino acid and the other containing the amino acid. Cells are grown for 3 days at 30°C and 8 colonies are restreaked on a fresh plate and grown for another 24hrs before being tested for integration.

10x TE buffer
pH 7.5
100mM Tris-HCl
10mM EDTA
Sterile filter; store at RT

1M Lithium-Acetate
dissolve 20.4g LiAc-dihydrate in H_2O to 200ml end-volume
Sterile filter; store at RT

LiAc/TE/PEG
100µl 1M LiAc
100µl 10x TE buffer
800µl 50% PEG
make fresh before use

50% PEG4000
50g PEG4000
dissolve in 100ml H_2O
Sterile filter; store at RT

LiAc/TE buffer
1ml 1M LiAc
1ml 10x TE buffer
8ml H_2O
make fresh before use

2.3.2 Integration PCR

To check if the plasmid has integrated correctly, that is at the right locus, not as a tandem repeat and the full plasmid rather than only the selective marker, a yeast-specific colony PCR is performed. Primers are designed to anneal just before and after the integration site in the genome and in the vector before and after the selective marker. If both these fragments are successfully amplified, the plasmid has positively integrated. To exclude clones with tandem integration, a third PCR reaction is set up, using the two primers on the vector. A PCR product is only possible, if there is a tandem repeat of the vector in the genome, and those clones are excluded.

Integration PCR		**PCR program**		
10x MangoTaq buffer	5µl	94°C	5min	
$MgCl_2$ (10mM)	1µl	94°C	30sec	
dNTPs (10mM)	1.5µl	54°C	30sec	35x
fw primer (10µM)	1µl	72°C	1.5min	
rv primer (10µM)	1µl	72°C	5min	
Template*	1µl	4°C	∞	
MangoTaq polymerase	0.5µl			
H_2O	14µl			
Total volume	25µl			

* As yeast cells are hardier than bacteria, they cannot be added straight into the PCR mix. Instead a small amount of cells are resuspended in 20μl 15mM NaOH and boiled for 5-10min. Cell debris is then removed by a short centrifugation at 3000g and 1μl of supernatant is used as a template for each PCR reaction.

2.3.3 Expression

Positive clones, without tandem integration, are re-streaked on the appropriate selective plate, to avoid carrying over of non-transformed yeast cells. This is important as the expression culture is grown in YP-media, rather than selective media, and therefore, also non-transformed cells can multiply. Similarly, as the selective pressure is lifted, it is possible for the integrated plasmid to be lost. Therefore, it was attempted to inoculate the expression culture with a large number of cells, grown on a selective plate overnight, to reduce the time the cells grow in YP-media. As selective plates contain glucose as a carbon source and genes under the Gal1-10 promoter are fully repressed under this growth condition, it can also be beneficial, especially if the recombinant expressed genes are slightly toxic for the cells. For a 5-litre expression, cells from one plate are collected and resuspended in a small amount of YP-media. 2.5-litre of YP-media supplemented with 2% raffinose is inoculated with these cells in the evening and grown overnight at 30°C, 220rpm. Cells should be at an OD_{600} of around 2 the next morning. Every culture is diluted with the same amount of YP-media and expression is induced by adding galactose to a final concentration of 2%. As yeast cells consume more oxygen than bacteria, only 500ml of culture is grown in each of the 2-litre conical flask. Protein is expressed for 4hrs and cells are subsequently harvested by centrifugation at 3000g, 4°C, 10min. The cell pellets are washed once with lysis buffer and pelleted again via centrifugation. An equal amount of lysis buffer, supplemented with EDTA-free protease inhibitor cocktail set III (Millipore) is added and cells resuspended. The suspension is frozen, by adding it drop-by-drop into LN_2 to produce the so-called ‘popcorn’, which is ready for cell lysis, or can be stored at -80°C.

2.3.4 Cell lysis

Yeast cells are more difficult to lyse than bacteria due to their robust cell wall. Sonication is only successful if the cell wall is enzymatically removed beforehand. This, however, is only feasible for a small amount of cells, whereas mechanical lysis is commonly used for large scale yeast cultures. Here, cells were mechanically lysed with a SPEX freezer mill 6875D (AXT). Popcorn, as well as a cooled, metal impactor, is added into the precooled freezer mill containers. The freezer mill, cooled to LN_2 temperature, rapidly shakes the container to crunch the popcorn and break the cells. After six cycles of shaking and pausing to avoid overheating, the popcorn has converted to cell powder, which can either be stored at -80°C or used directly.

2.4 Baculovirus/insect cell protein expression

Over the last three decades insect cells have gained more and more importance as an expression system for eukaryotic proteins and protein complexes. A genetically engineered baculovirus carrying the genes of interest is used to infect the cells and therefore, induce recombinant protein expression. Insect cells represent a sophisticated expression system and are capable of all PTMs, while at the same time are still easier to handle and to grow to a high cell number than mammalian cells.

2.4.1 MultiBac system

The MultiBac system (Fitzgerald et al., 2006) is specifically designed for multiprotein complex expression in insect cells. Each transfer vector has two MCSs under the control of the two baculoviral promoters. Multiple vectors can be combined, either by transfer of expression cassettes or by *in vitro* Cre-fusion, allowing for flexibility of testing multiple gene combinations. A bacmid is produced from the vectors, facilitated by specialised cells, and purified bacmid can be used to generate the virus. This virus can infect multiple large-scale insect cell cultures for protein expression.
Codon-optimised genes of all four CBF3 subunits were cloned into the two MCSs of a donor and an acceptor vector, as described in 2.1.2, with the only exception being that the donor vector has to be transformed in PIR1 rather than XL1 cells, as it carries a different origin of replication. Insertion was verified via sequencing and the two

vectors were combined through incubation with Cre-recombinase for 1hr at 37°C and subsequent heat inactivation for 10min at 70°C. 5μl was transformed into XL1 cells and a miniprep is made from one single clone.

Cre-recombination:

Donor vector + genes	500ng
Acceptor vector + genes	500ng
Cre buffer (10x)	1μl
Cre recombinase	1μl
H_2O	xμl*
Total volume	10μl

* appropriate volume to make up to total volume

Transposition is conducted in DH10MultiBac cells. 5μl of recombined plasmid is transformed in 50μl of cells, which are subsequently grown in 900μl of SOC medium for 5hrs. 30μl of cells are plated on LB-Agar plates containing the appropriate antibiotics, 0.5mM IPTG and 42μg/ml XGAL. The last two components are added to allow for blue/white screening, as the Tn7 transposition element is embedded in the LacZ gene. Therefore, positively transpositioned cells lose this gene and fail to produce blue colouring. For bacmid preparation a ONC is grown from a white, positive colony, and cells are pelleted via centrifugation. Cells are lysed using 250μl P1 buffer from a miniprep kit (Qiagen). Cell debris and proteins are precipitated by addition of 250μl P2 and 350μl N3 buffer and removed by centrifugation at 13,000g for 10min at RT. All subsequent steps are performed in a tissue culture hood and all solutions used are sterile filtered to avoid any contaminations. The cleared supernatant is transferred into a sterile 1.5ml Eppendorf tube and again centrifuged to ensure no cell debris is carried over. 800μl isopropanol is added and the mixture is incubated on ice for 10min. DNA is pelleted at 13,000g, for 15min at 4°C, and supernatant carefully removed. The bacmid DNA pellet is air-dried in the tissue culture hood and subsequently dissolved in 30μl H_2O and can be stored at -20°C or used immediately.

Transfection, to produce virus is carried out in Sf21 cells in 6-well plates. 2ml of $5x10^5$ cells in antibiotic-free media are added per well and left to settle for one hour. In the

meantime, 5μl of bacmid is mixed with 100μl of antibiotic-free media and 10μl of Gene Juice transfection reagent (Novagen) is diluted in 100μl of antibiotic-free media, before both are added together and incubated for 45min. After this incubation 800μl antibiotic-free media is added to the now ready to use transfection mix. The media in the 6-well plates is carefully withdrawn and replaced immediately with 1ml of the transfection mix. Cells are incubated for 5-6hrs at 28°C. Transfection mix is carefully withdrawn and 2ml of media supplemented with antibiotics is added dropwise to the cells. Cells are incubated at 28°C for 72hrs. After this, cells are resuspended and added to 50ml of fresh Sf21 cells, diluted to a cell density of $1.8x10^6$. This culture is grown for 3 days or until cell viability has dropped to 70-80% due to virus production. Cells are harvested via centrifugation at 1000g, 4°C for 10min and the media, containing the virus, is carefully transferred into a new black falcon tube and a final concentration of 2% FBS is added. Virus can then be stored in the cold room or frozen.

2.4.2 Infection and protein expression

Fresh virus needs to be produced before every large-scale infection for protein expression. Therefore, 100ml of $1.8x10^6$ Sf21 cells are infected with 50μl of virus stock and grown for 3 days until cell viability has dropped to 70-80%. Cells are harvested via centrifugation at 1000g, 4°C for 10min and the supernatant is used to infect a large-scale culture of HiFi cells, diluted to a cell density of $2 x10^6$ in media supplemented with antibiotics. Infected cells are grown 3 days to allow for protein expression. Cells are harvested by centrifugation, and cell pellet can be shock-frozen or used immediately.

2.4.3 Cell lysis

Pellet is resuspended in 50ml lysis buffer with added EDTA-free protease inhibitor cocktail set III (Millipore) and 5μl Benzonase is added to digest insect cell DNA. Cells are lysed through sonification as described in 2.2.3. Lysate is centrifuged at 30,000g for 50min to remove cell debris and supernatant is filtrated with a 5μm syringe filter, before being used immediately for purification.

2.5 Protein purification

Typical recombinant protein purification uses a number of characteristics of the protein of interest to separate it from the cell lysate. Usually, as a first step affinity purification is performed, using a N- or C-terminal tag, which was fused to the protein of interest for this exact purpose. Depending on purity, a number of other purification steps can be performed, including another affinity purification step, anion/cation exchange chromatography, which exploits the charge of the protein of interest, or size exclusion chromatography, using size differences to separate proteins. The latter also provides an insight on quality of the protein sample, as aggregation or a mixed oligomeric state can be detected, as well as protein complex formation can be judged. To assess purity, samples are separated via SDS-PAGE, and chromatographies are run using an ÄKTA protein purification system (GE Healthcare), which allows to track protein and DNA contaminations through UV measurements at 280nm and 254nm respectively. Every purification protocol has to be optimised for the protein of interest, and protocols used in this thesis are described in detail below.

2.5.1 Proteins expressed in bacterial

Histones: In contrast to human, all four budding yeast histones can be co-expressed and purified as octamers for the reconstitution of nucleosomes or other uses (Kingston et al., 2011). This is possible with the canonical (H3, H4, H2A and H2B), as well as the centromeric (Cse4, H4, H2A and H2B) octamer. As over-expression of histones is toxic to the bacteria cells, expression was carried out at 37°C for 4hrs. Nonetheless, expression levels are reasonably high and 2-4litre culture per purification was sufficient. Histones can be purified untagged via a 5ml HiTrap Heparin HP (GE Healthcare) column, followed by a size-exclusion chromatography to separate DNA from the octamer. Cells were lysed in lysis buffer, loaded onto the Heparin column, which was subsequently washed with buffer A containing 0.5M salt to wash of most impurities. Histones were then eluted with a gradient from 0.5M salt (buffer A) to 2M salt (buffer B). Histone octamer bind most stably to the Heparin column and therefore elute last. The octamer is most resilient after this purification step and can be shock frozen, or directly further purified. To separate the octamer from residual *E.coli* DNA, it is subjected to size exclusion chromatography (Superdex

200, 10/300 GL; GE Healthcare) in 2M salt (GF buffer). After this last step, the pure octamer is concentrated via a 5kD Vivaspin 20 concentrator (Sartorius) to a final concentration of 7-8μM and must be used immediately, due to its instability without DNA present. The same procedure was used to purify the octamer containing the canonical H3, as well as the centromeric full-length Cse4 or Cse4 HFD, which only contains the histone fold domain, but not the long, non-conserved N-terminal tail.

Lysis Buffer	**Buffer A**	**Buffer B/GF**
20mM HEPES, pH 8	20mM HEPES, pH 8	20mM HEPES, pH 8
100mM NaCl	500mM NaCl	2M NaCl
1mM DTT	1mM DTT	1mM DTT
0.1mM EDTA	0.1mM EDTA	0.1mM EDTA

Mif2 constructs: Two constructs of Mif2 were expressed and purified: Mif2-530, as used in Xiao et al., 2017, which has a N- and C-terminal deletion, and Mif2-end which only has the N-terminal deletion but includes the last 19 amino acids (see Table 8 for details). Both constructs were purified with the same procedure, using a N-terminal his tag for affinity purification with a 5ml HisTrap HP (GE Healthcare) as the first step. Therefore, lysate was loaded onto the column and washed with lysis buffer until the UV baseline was steady. As Mif2 binds strongly to DNA, the column was washed with 3CV salt wash buffer, before protein was eluted with a shallow gradient from lysis buffer to buffer B, which contains 500mM imidazole. Collected fractions are run on SDS-PAGE to identify those containing the Mif2 protein. Fractions are pooled and subjected to anion exchange chromatography with a 1ml Poros 50 HQ (Thermo Fisher) column. Protein is loaded onto the column, washed with buffer C, and eluted with a salt gradient from buffer C to buffer D. Fractions are again tested with SDS-PAGE and those containing the Mif2 protein are pooled. Anion exchange is particularly useful to separate any remaining DNA contaminations, as DNA binds very strongly to the anion exchange resin and elutes after the protein. Pooled protein is concentrated using a 30kD Vivaspin 20 concentrator (Sartorius), and subjected to size exclusion chromatography using a Superdex 200 10/300 GL (GE Healthcare) column in GF buffer. Notably, size-exclusion showed a salt dependency on protein stability/aggregation of both Mif2 constructs, exhibiting a good elution profile in 300mM but not in 150mM salt. Protein is concentrated using a 30kD Vivaspin 20

concentrator (Sartorius), 4% glycerol is added and it is shock-frozen in LN_2 and stored at -80°C.

Lysis buffer	**Salt wash buffer**	**Buffer B**
50mM HEPES, pH 7.5	50mM HEPES, pH 7.5	30mM HEPES, pH 7.5
150mM NaCl	1M NaCl	150mM NaCl
1mM TCEP	1mM TCEP	1mM TCEP
10mM Imidazole	2.5% glycerol	500mM Imidazole
5% glycerol		

Buffer C	**Buffer D**	**GF buffer**
20mM HEPES, pH 7.5	20mM HEPES, pH 7.5	10mM HEPES, pH 7.5
120mM NaCl	1M NaCl	300mM NaCl
1mM DTT	1mM DTT	1mM DTT

Ndc10 constructs: Ndc10 constructs expressed in bacteria included the NTD, CTD, CTDshort and centralCTD (see Table 8). NTD, for which the crystal structure was determined, was already expressed and purified by a former lab member (Perriches, 2014), and pure frozen protein was used for assays. All other constructs were expressed and purified. CTD had a N-terminal his tag followed by a sumo tag, for affinity purification and improved solubility respectively. CTDshort and centralCTD contained a N-terminal his-tag for affinity purification. Lysate was loaded onto a 5ml HisTrap HP (GE Healthcare) column and was subsequently washed with lysis buffer, until the UV baseline was steady. Protein was eluted with an imidazole gradient from lysis buffer to buffer B. SDS-PAGE was used to assess which fraction contains the Ndc10 constructs and these fractions were pooled. His-tag was cleaved by adding a 1:100 dilution of 1mg/ml Sumo or TEV protease and incubation overnight at 4°C. Anion exchange chromatography using a 1ml Poros 50 HQ (Thermo Fisher) or Mono Q 5/50 GL (GE Healthcare) column was used to separate the Sumo or TEV protease, cleaved tag and other protein and DNA contaminations. Protein was diluted with NS buffer to a final salt concentration of 150mM, subsequently loaded onto the column, which was washed with buffer C and eluted with a salt gradient from buffer C to D. Fractions containing the protein were pooled and concentrated with a 10kD Vivaspin 20 concentrator (Sartorius). Lastly, size exclusion chromatography was performed, using a HiLoad 16/60 Superdex 200 (for CTD) or 75 (for the other two constructs;

GE Healthcare) in GF buffer. Protein fractions were pooled, concentrated and either used immediately or shock-frozen in LN_2 with 4% glycerol.

Lysis buffer	**Buffer B**	**Buffer C**
50mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8
400mM NaCl	200mM NaCl	150mM NaCl
1mM TCEP	1mM TCEP	1mM TCEP
10mM Imidazole	500mM Imidazole	
5% glycerol		

Buffer D	**GF buffer**	**NS buffer**
20mM HEPES, pH 7.5	20mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8
1M NaCl	150mM NaCl	1mM TCEP
1mM TCEP	1mM TCEP	

Scm3: N-terminal his-tagged Scm3 was expressed for interaction studies with Ndc10. Lysate was loaded onto an equilibrated 5ml HisTrap HP (GE Healthcare) column and washed until the UV baseline became steady. Bound protein was eluted with an imidazole gradient from lysis buffer to buffer B and fractions were run on a SDS-PAGE. Fractions with Scm3 protein were pooled and incubated at 4°C, overnight with 0.5ml TEV protease (1mg/ml) to remove the his-tag. Protein was diluted with an appropriate volume of NS buffer to reduce the salt concentration to 100mM. It was subsequently loaded on a Mono Q 5/50 GL (GE Healthcare) column, washed with buffer C and eluted with a salt gradient from buffer C to D. This step was crucial to remove DNA impurities. Protein peak fractions were pooled and concentrated to 5ml with a 10kD Vivaspin 20 concentrator (Sartorius) and subsequently run on a HiLoad 16/60 Superdex 75 (GE Healthcare) in GF buffer. Protein quality was tested with SDS-PAGE, protein was pooled and used either immediately or shock-frozen in LN_2 with 4% glycerol.

Lysis buffer	**Buffer B**	**Buffer C**
50mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8
150mM NaCl	150mM NaCl	100mM NaCl
1mM TCEP	1mM TCEP	1mM TCEP
10mM Imidazole	500mM Imidazole	

Buffer D	**GF buffer**	**NS buffer**
20mM HEPES, pH 7.5	20mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8
1M NaCl	150mM NaCl	1mM TCEP
1mM TCEP	1mM TCEP	

Scm3 NTD and Ndc10 CTD co-expression: N-terminal his-tagged Scm3 NTD and Ndc10 CTD (see Table 8 for details) were co-expressed in bacteria and cells were lysed in lysis buffer. Proteins were co-purified with a 5ml HisTrap HP (GE Healthcare) column by loading the lysate and washing the column with lysis buffer until UV baseline was steady. Proteins were eluted with an imidazole gradient from lysis buffer to buffer B. The his-tag from Scm3 was cut at 16°C for 2hrs at 120rpm. Protein was diluted to 150mM salt concentration with NS buffer, to be able to load it onto a Mono Q 5/50 GL (GE Healthcare) column, equilibrated with buffer C. Non-bound protein was washed off the column with buffer C, and bound fraction was eluted with a salt gradient from buffer C to D. This step is particularly useful to separate excess Scm3 and also partly the Ndc10 CTDshort degradation, as well as DNA impurities. Peak fractions were pooled and concentrated with a 5kDa Vivaspin 20 concentrator (Sartorius). Size exclusion chromatography with a HiLoad 16/60 Superdex 75 (GE Healthcare) in GF buffer was performed to separate impurities, as well as Scm3 NTD/Ndc10 CTD complex from individual proteins. Fractions were tested with SDS-PAGE, pooled and concentrated. Concentrated protein was either immediately used or flash-frozen with 4% glycerol.

Lysis buffer	**Buffer B**	**Buffer C**
50mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8
300mM NaCl	300mM NaCl	150mM NaCl
1mM TCEP	1mM TCEP	1mM TCEP
10mM Imidazole	500mM Imidazole	

Buffer D	**GF buffer**	**NS buffer**
20mM Tris-HCl, pH 8	10mM Tris-HCl, pH 8	20mM Tris-HCl, pH 8
1M NaCl	200mM NaCl	1mM TCEP
1mM TCEP	1mM TCEP	

Cse4/H4/Scm3: This complex can be co-expressed and purified with a N-terminal his-tag on Cse4 via affinity purification with a 5ml HisTrap HP (GE Healthcare) column. Fractions containing all three proteins are pooled and tag is cleaved overnight at 4°C with addition of 250µl TEV protease (1mg/ml). Protein solution is diluted with NS buffer to ~350mM salt concentration. To separate impurities, as well as DNA contaminations, anion exchange chromatography with a Mono Q 5/50 GL (GE Healthcare) or 1ml Poros 50 HQ (Thermo Fisher) column is performed. Fractions are tested with SDS-PAGE and these containing all three proteins are pooled, concentrated and flash-frozen for further use. It should be noted that, whilst co-expression and purification is possible, proteins are not very stable and the yield is limiting.

Lysis buffer
50mM HEPES, pH 7.5
500mM NaCl
1mM DTT
20mM Imidazole
10% Glycerol

Buffer B
20mM HEPES, pH 7.5
300mM NaCl
1mM DTT
500mM Imidazole

Buffer C
20mM HEPES, pH 7.5
350mM NaCl
1mM DTT

Buffer D
20mM HEPES, pH 7.5
1M NaCl
1mM DTT

NS buffer
20mM HEPES, pH 7.5
1mM DTT

2.5.2 Proteins expressed in yeast

CBF3, CBF3 core, and all other CBF3 constructs: All CBF3 constructs used in this thesis (see Table 8) have been purified with the same procedure described here. All four CBF3 genes, as well as Gal4 and Sgt1, are co-expressed and processed to cell powder. Cell powder is dissolved in twice the volume of lysis buffer (e.g. 50ml falcon of cell powder in 100ml lysis buffer), supplemented with protease inhibitor cocktail set III (Millipore). The addition of phosphatase inhibitor (50mM NaF) to the lysis buffer is essential for the purification of the whole complex (i.e. not lacking Ndc10) but also to obtain a good yield of both core and full complex. After the cell powder is fully dissolved, lysate is briefly sonicated with a Branson Sonifier 450 at step 2, four times

for 10sec on wet ice, to shear DNA. To remove insoluble proteins and cell debris, it is subsequently centrifuged at 30,000g for 50min at 4°C. StrepTrap HP (GE Healthcare) column is used as a first purification step. Therefore, Cep3 was tagged on the C-terminus, which seemed to be the only position to enable full complex purification. The column is washed with water, regenerated with 2CV 0.5M NaOH, before being washed with water and equilibrated with lysis buffer. Lysate is loaded at 2.5ml/min, and the column is washed with lysis buffer, supplemented with protease inhibitor cocktail set III (Millipore), until the baseline is just steady. As Ndc10 is not very stably bound, it is not advisable to wash longer. Protein is eluted with 30ml lysis buffer supplemented with 20mg Desthiobiotin and the peak fraction is pooled and loaded onto a 5ml HiTrap Heparin HP (GE Healthcare) column immediately. Heparin purification is performed to separate a prominent high molecular weight impurity, which otherwise co-elutes on size exclusion chromatography, due to a similar size. It also offers the advantage of separating excess Cep3 (especially prominent without co-expression of Sgt1) and Ndc10 from the core complex (see next paragraph for more details). SDS-PAGE is used to assess which fraction contain core and Ndc10, and said fraction are pooled together. To purify only the core complex, fractions containing Ndc10 can be omitted at this step. 250μg TEV protease is added to cut the tag overnight at 4°C. The complex is further purified by 1ml Poros 50 HQ (Thermo Fisher) column. This step allows for the removal of residual DNA, as well as a second prominent impurity around 100kDa and 50kDa, identified as glutaminyl- and methionyl-tRNA synthetase and their cofactor ARC1 by mass spectrometry. Alternatively, Mono Q 5/50 GL (GE Healthcare) can be performed, which offers the additional advantage in separating the core from the full complex, if eluted with a shallow gradient. CBF3 proteins are pooled and concentrated using a 100kDa Vivaspin 20 concentrator (Sartorius) at 4000g, 4°C and subsequently, size exclusion chromatography with a Superose 6 10/300 GL (GE Healthcare) is performed to separate aggregations from the full and core complex. Salt concentration is kept at 300mM, as both complexes are salt sensitive and buffers with lower salt show extensive aggregation. For structural studies protein is used immediately after size exclusion chromatography, for assays it was concentrated if necessary and shock frozen with 4% glycerol.

Lysis buffer:	**Buffer B:**	**Buffer C:**
50mM HEPES, pH7.5	30mM HEPES, pH 7.5	20mM HEPES, pH 7.5
250mM NaCl	1M NaCl	200mM NaCl
1mM DTT	1mM DTT	1mM DTT
50mM NaF	10mM NaF	5mM NaF
10% glycerol	5% Glycerol	
0.02% NP-40	0.02% NP-40	

Buffer D:	**NS Buffer:**	**GF Buffer:**
20mM HEPES, pH 7.5	30mM HEPES, pH 7.5	10mM HEPES, pH 7.5
1M NaCl	1mM DTT	300mM NaCl
1mM DTT	10mM NaF	1mM DTT
2.5mM NaF		

Ndc10 FL and Cep3: Untagged Ndc10 full-length and tagged Cep3 full-length, used for assays, were derived through expression of the full complex and separated during purification as described above. For Cep3 it was crucial to express the CBF3 proteins without simultaneous co-expression of Sgt1, as this yields a high percentage of excess Cep3. The Heparin step is used to separate both Ndc10 and Cep3 from the core complex. Excess Cep3 does not bind to the 5ml HiTrap Heparin HP (GE Healthcare) column and can therefore be collected in the flow-through. Ndc10 binds strongest to the Heparin column and can be separated during elution with a shallow salt gradient. Both proteins must be pooled and further purified as soon as possible after the Heparin run, as they are inherently unstable and degrade rapidly. To separate protein and DNA contamination, anion exchange chromatography with a 1ml Poros 50 HQ (Thermo Fisher) column is performed and protein is pooled and subsequently concentrated as soon as the peak elutes. Protein is shock-frozen with 4% glycerol at the required concentration and stored at -80C.

Buffers – same as CBF3 purification

Skp1/Ctf13: Co-expressed N-terminal tagged Skp1 and untagged Ctf13 (with additional co-expression of Gal4 and Sgt1) was purified via a similar protocol as the full CBF3 complex. StrepTrap HP (GE Healthcare) column was used as initial affinity chromatography step. Lysate was loaded and column washed with lysis buffer until the UV baseline was steady. Although proteins eluted very pure, an anion exchange

chromatography with a 1ml Poros 50 HQ (Thermo Fisher) column was performed to further separate contaminations. Therefore, protein was diluted with NS buffer to a salt concentration of 200mM, loaded and eluted with a salt gradient from buffer C to D. Protein was pooled and concentrated with a 30kDa Vivaspin 20 concentrator (Sartorius) and subsequently run on a Superdex 200 10/300 GL (GE Healthcare) column in GF buffer. Here it should be noted, that similar to the full CBF3 or core complex Skp1/Ctf13 is sensitive to salt concentration and was therefore kept in a buffer containing 400mM salt. Size exclusion in a buffer with 150mM salt leads to complete aggregation, whereas the elution profile with 300mM salt shows some aggregation, a highly asymmetric peak, as well as disruption of the two proteins. Protein was pooled, concentrated and either used immediately or shock-frozen with 4% glycerol.

Buffers – same as CBF3 purification, except GF buffer
GF Buffer:
10mM HEPES, pH 7.5
400mM NaCl
1mM DTT

2.5.3 Proteins expressed in insect cells

CBF3: CBF3 expressed in insect cells was subjected to affinity purification with a 5ml StrepTrap HP (GE Healthcare) column. Lysate was loaded onto the column and washed with lysis buffer until the baseline was steady. Bound protein was eluted using 30ml of lysis buffer complemented with 0.7mg/ml desthiobiotin. Collected fractions were tested for expression with SDS-PAGE.

Lysis buffer:
50mM HEPES, pH7.5
150mM NaCl
1mM DTT
5mM EDTA
5mM NaF
7% glycerol
0.02% NP-40

2.6 Nucleosome reconstitution

Nucleosomes can be reconstituted *in vitro* from histone octamers and DNA fragments. Generally, one would use a strong positioning DNA sequence, such as the 147bp Widom 601 sequence (Lowary and Widom, 1998) to enhance nucleosome formation. As CBF3 binds sequence-specifically to the centromere, however, nucleosomes were reconstructed with both the centromeric sequence and the Widom 601 sequence (see Table 9 for exact sequences used). Histone octamers and DNA are mixed in 2M salt buffer and salt concentration is slowly decreased to force spontaneous nucleosomes formation.

2.6.1 DNA-Preparation

Two common methods are used to produce enough DNA for nucleosome refolding purpose. DNA is either cloned into a bacterial plasmid with flanking restriction sites, plasmid DNA is purified in large scale and the desired DNA fragment is cut out of the plasmid and purified, or DNA can be amplified via PCR on a large scale and then purified. The latter was used in this thesis, as it also allows for easy tagging of the DNA by the use of biotinylated primers. Both budding yeast centromere DNA and Widom 601 DNA was amplified in large scale with appropriate primers, in-house purified Pfu polymerase (5mg/ml) and optimised PCR parameters. PCR quality is checked via a 2% agarose gel electrophoresis and PCR product is pooled and filtered with a 0.2μm syringe filter.

PCR (for four 96-well plates)	
1M Tris-HCl, pH 8	800μl
3M KCl	133μl
1M $MgCl_2$	80μl
10% Triton	400μl
H_2O	37.4ml
→ mix well	
20mg/ml BSA (NEB)	200μl
dNTPs (25mM)	400μl
fw primer (500μM)	80μl
rv primer (500μM)	80μl
template (250ng/μl)	48μl
Pfu polymerase	10μl
Total volume	40ml

PCR program

98°C	2min	
98°C	30sec	
53°C	30sec	30x
72°C	30sec	
72°C	2min	
4°C	∞	

DNA is precipitated at -20°C overnight by adding 3M NaAcetate to a final concentration of 10% and 2.5x ice-cold 100% EtOH. DNA is pelleted via centrifugation at 5000g, 4°C for 30min. Supernatant is carefully decanted and DNA pellet is resuspended in 50-100ml buffer A and loaded onto an equilibrated Poros 50 HQ (Thermo Fisher Scientific) column. Column is washed with 20CV buffer A, followed by a wash with 22% buffer B to elute residual dNTPs and primers. DNA is then eluted with a gradient from 22% to 45% buffer B in 15CV. Collected fractions are tested on an agarose gel, DNA is pooled accordingly and concentration is measured with a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific). DNA is again precipitated as described above, pelleted via centrifugation and the DNA pellet is dissolved in H_2O to a final concentration of 3-4mg/ml.

Buffer A	**Buffer B**
10mM Tris-HCl, pH 8	10mM Tris-HCl, pH 8
1mM EDTA	1mM EDTA
	2M NaCl

2.6.2 Nucleosome formation

Nucleosomes can spontaneously form *in vitro*. Therefore, both histone octamers and DNA, prepared in a buffer containing 2M salt, are slowly dialysed to reduce the salt concentration. As mentioned before, this is most efficient with a strong positioning sequence, but also possible with other DNA sequences, like the centromere.
Prepared DNA and Histone octamers, expressed and purified as described in 2.5.1, are mixed in a molar ratio of 1:1 in a final volume of 100µl. Salt concentration is kept at 2M salt and a final concentration of 10mM DTT is added. The mix is pipetted carefully in a 2K Slide-A-Lyzer MINI Dialysis Device (Thermo Fisher Scientific) and dialysed against 700ml 2M salt buffer for 30-60min. Salt concentration can then be lowered either gradually by pumping a buffer containing no salt (NS buffer) into the dialysis container, or by stepwise dialysis. Both methods were used in this thesis, and no difference in results was noticed. The former, uses a peristaltic pump at around 3.5ml/min flow rate and dialysis was performed overnight, followed by dialysis against 700ml of final buffer for 1hr. For the latter, dialysis is performed against 700ml of 1M salt buffer overnight, with subsequent dialysis for 2-3hrs each against 700ml

of 0.8M, 0.66M, 0.3M and 0.2M buffer, followed by a last dialysis step against the final buffer overnight. Both dialyses were performed at 4°C and nucleosomes were carefully pipetted from the dialysis unit and into a fresh Eppendorf. To remove any precipitation sample was centrifuged at 4000rpm, 4°C for 2min.

Dialysis buffers
10mM HEPES, pH7.5
2M/1M/0.8M/0.66M/0.3M/0.2M/0.12M (**final buffer**) or 0M (**NS buffer**)
1mM DTT

2.6.3 Native-PAGE

To ensure nucleosome formation, samples can be run on a native PAGE. This will separate any remaining free DNA and nucleosomes, whereas the latter migrates higher due to its bigger size. 10µl nucleosome sample is mixed with 2µl DNA-loading buffer and run on a 0.8% agarose gel for 20min at 80V. Both buffer and gel are precooled and placed on ice during the run to prevent heating of the gel/sample. As control a sample containing only free DNA is run on the same gel. To visualise DNA, gel is incubated for 20-30min with 50ml H_2O containing 2µl EtBr and subsequently imaged with UV light in a Gel Doc™ EZ Imager (BioRad). Afterwards, to visualise proteins, the gel is stained with InstantBlue (Expedeon). Alternatively, a 6% DNA retardation gel (Invitrogen) can be used, and run with precooled 1xTBE buffer for 40min at 220V.

2.7 Protein analysis

2.7.1 SDS-PAGE

SDS-PAGE (Sodium dodecyl sulfate – polyacrylamide gel electrophoresis) is an analytical method to separate proteins according to their size in an electric field. Both commercial available and home-made gels were used. The latter were prepared following the recipe shown below and kept in the fridge for up to one week. 12.5µl protein sample was mixed with 5µl protein loading buffer and heated at 98°C for 2min before loading. Unless stated otherwise, 2.5µl of PageRuler™ Plus Prestained Protein Ladder (GE Healthcare Life Sciences) was used as a marker. Gels were run

at 220V for 45min in SDS-running buffer (homemade gels) or 1x MOPS running buffer (commercial gels), before being stained with InstantBlue (Expedeon).

10% Separating gel
375mM Tris-HCl (pH 8.8)
10% polyacrylamide (37.5:1)
0.1% SDS
0.05% TEMED
0.1% APS

15% Separating gel
375mM Tris-HCl (pH 8.8)
10% polyacrylamide (37.5:1)
0.1% SDS
0.05% TEMED
0.1% APS

4% Stacking gel
125mM Tris-HCl (pH 6.8)
4% polyacrylamide (37.5:1)
0.1% SDS
0.1% TEMED
0.1% APS
~0.05% Bromophenol blue

2.7.2 Protein concentration measurement

Protein concentration was measured with a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific). For pure protein, both molecular weight and extinction coefficients were provided for a more accurate calculation of concentration.

2.7.3 SEC-MALS

Size exclusion coupled with multi-angle light scattering is used to determine the molecular mass and therefore oligomeric state of proteins and protein complexes. 200μg of either core or full CBF3 complex were injected onto a Superose 6 Increase 10/300 (GE Healthcare) column, protein is separated due to its mass and injected into a Dawn 8+ MALS system (Wyatt). Data was analysed using Astra software.

2.7.4 AUC

Analytical ultra-centrifugation can be used to determine the mass of a protein and therefore the oligomeric state, by measuring the sedimentation velocity of a protein in solution. It can also be used to test interaction between different proteins,

especially if this interaction is weak and cannot withstand more stringent methods like size exclusion chromatography. In principle, an AUC is a normal ultra-centrifuge fitted with a high-speed camera and a special rotor which features an optical window to observe the sample during the run.
AUC was conducted with the Ndc10 CTDshort construct to measure its oligomeric state and test if this construct likely aggregates. Measurements were made in low salt buffer with 0.9, 0.6 and 0.3mg/ml, as well as in high salt buffer with 1, 0.8, 0.4 and 0.2mg/ml and centrifugation at 35,000rpm, at 16°C for 20hrs.

Low salt buffer:
20mM Tris-HCl, pH 7.5
150mM NaCl
1mM TCEP

High salt buffer:
20mM Tris-HCl, pH 7.5
400mM NaCl
1mM TCEP

2.7.5 Phos-tag SDS-PAGE

Phos-tag reagent is a simple and effective way to identify phosphorylated proteins via SDS-PAGE. The reagent is included into the gel and interacts with phospho-groups. Therefore, phosphorylated proteins exhibit a shift towards higher molecular weight. The reagent was used accordingly to manufacture's recommendation with a 12% polyacrylamide SDS gels (see 2.7.1 for details). During the casting 100μM Phos-tag and 200μM $MnCl_2$ were included into the separating gel. Untreated and phosphatase treated samples with added 3μl of SDS loading buffer, were boiled and 10μl was run at 220V for 45min in SDS-running buffer. Gels were stained with InstantBlue (Expedeon).

Phosphatase reaction:

Core (0.3mg/ml)	8μl
λ phosphatase	0.5μl
H_2O	1.5μl
Total volume	10μl

2.7.6 Western blot

Western blot coupled to SDS-PAGE allows for detection of specific proteins in a heterogeneous sample, even if the protein of interest is only present in very small amounts. Antibodies are used for detection, usually in a combination of primary, which is specific against the protein or protein tag, and secondary antibody, which is specific against the former and allows for detection via immunofluorescence or radioactivity. Denatured proteins are separated via SDS-PAGE and transferred from the gel to a nitrocellulose membrane using the iBlot (Invitrogen). Protein bands on the membrane are temporarily stained with Ponceau S solution to check on the quality of the transfer. The membrane is destained with water, and washed with PBS supplemented with 0.1% Tween-20 (PBS-Tween), before being blocked for 30min with 5% milk in PBS-Tween on a shaking platform at RT. The membrane is washed once with PBS-Tween and incubated with HA-tag antibody, diluted 1:5000 in PBS-Tween and 2% milk for 1hr at RT, shaking. Primary antibody is decanted and can be stored for a short time at 4°C, or for a longer period at -20°C and reused up to three times. The membrane is washed three times with PBS-Tween for 5min at RT, shaking and is then incubated with secondary anti-mouse HRP antibody (GE Healthcare) for detection for 1hr at RT, shaking. Secondary antibody is decanted and membrane is again washed for three times as described above. Shortly before detection, the membrane is placed on a clear film and 1ml of freshly mixed Amersham ECL™ Western Blotting Detection reagents (GE Healthcare) solution is added. Membrane is placed in an ImageQuant™ LAS 4000 imager (GE Healthcare) and exposed for 30sec to 5min depending on signal strength.

2.8 Protein/protein and protein/DNA interaction studies

2.8.1 Pull-down assay

A pull-down assay is an *in vitro* interaction assay, used to confirm known protein-protein interactions or find new ones. The principle is to use a tagged bait protein, which can be affinity purified using the appropriate resin, to capture prey protein(s), which are untagged, and can only be co-purified if they interact with the bait protein.

These assays can be done using co-expressed proteins and purifying straight from the lysate, or by using already purified proteins.

CBF3 mutant pull-down assay with co-expression in yeast: To test the influence of mutations in Skp1 (Skp1 NY_KK) and Cep3 (Cep3 SA or SE; see Table 8 for details) on complex formation, a pull-down assay with co-expressed proteins was conducted. As expression in yeast is not high enough to see recombinant proteins on SDS-PAGE of the lysate, all proteins, except for Ctf13, were HA-tagged. This enabled detection via HA-tag antibody and western blot. Ctf13 was not tagged as both N-terminal or C-terminal tagging disrupted complex formation, as tested with wild-type proteins. Co-expression of mutant and wild-type CBF3, as control, was done as described in 2.3.3. An equal amount of cell powder was used for both mutants and control and cell lysate was incubated with Streptavidin agarose beads (Invitrogen) for 20min. Beads were gently pelleted via centrifugation, supernatant was removed and beads were washed three times with lysis buffer. Beads were resuspended directly in SDS-loading buffer and proteins were visualised via SDS-PAGE and western blot with HA-tag antibody. In the case of the Skp1 NY_KK mutant, to enable the Coomassie visualisation of endogenous Skp1, a large-scale pull-down using a 5ml StrepTrap HP (GE Healthcare) column with co-expressed mutant and wild-type CBF3 was conducted. Eluate was separated via SDS-PAGE and both InstantBlue (Expedeon) stained and a western blot with HA-tag antibody was prepared.

Lysis buffer:
50mM HEPES, pH7.5
250mM NaCl
1mM DTT
50mM NaF
10% glycerol
0.02% NP-40

Ctf13/Skp1 pull-down from co-expressed proteins in yeast: To elucidate if two potential phosphorylation sites on Skp1 (S162 and T177) influence the binding of Ctf13, both proteins were co-expressed (see Table 8) with a 2xstrep-tag on the N-terminus of Skp1 to facilitate a pull-down. An equal amount of cell powder of co-expressed wild-type or mutant Skp1 and Ctf13 was dissolved in 60ml lysis buffer.

Lysate was loaded onto a 5ml StrepTrap HP (GE Healthcare) column, washed with 10CV lysis buffer and subsequently eluted with 30ml lysis buffer containing 20mg Desthiobiotin. Fractions were examined with 12% SDS-PAGE and Coomassie stain.

Lysis buffer:
50mM HEPES, pH7.5
250mM NaCl
1mM DTT
50mM NaF
10% glycerol
0.02% NP-40

CBF3 core formation pull-down assay with purified proteins: A pull-down assay with purified tagged Cep3 and untagged Skp1/Ctf13 (wild-type, Skp1Δ and Skp1 NY_KK; see Table 8) was performed to test *in vitro* complex formation. All three proteins were expressed and purified as described in 2.3.3 and 2.5.2. 50μl Cep3 (0.4mg/ml) and 35μl of the different Skp1/Ctf13 complexes (0.5mg/ml) were mixed and incubated for 5min on ice. 50μl Cep3 mixed with 35μl reaction buffer, and 35μl of the different Skp1/Ctf13 complexes mixed with 50μl reaction buffer, was used as controls. A 12.5μl aliquot was taken from each reaction to serve as an input sample for SDS-PAGE. 10μl Streptavidin agarose beads (Invitrogen), equilibrated with reaction buffer, was added to the rest of each reaction and binding of protein was facilitated by incubation for 1hr at 4°C on a shaking platform. Beads were pelleted by gentle centrifugation at 3000rpm for 20sec and washed three times with 150μl reaction buffer. Beads, as well as lysate aliquots, were mixed with 5μl SDS loading buffer, boiled for 5min and loaded on a 12% SDS-PAGE.

Reaction buffer:
10mM HEPES, pH7.5
300mM NaCl
1mM DTT

CBF3 formation pull-down assay with purified proteins: To elucidate CBF3 complex formation of the core with Ndc10 the same, above mentioned protocol was used. As input, 40μl of core (0.5mg/ml) was mixed with 132μl of Ndc10 full-length (0.3mg/ml), 50μl of Ndc10 NTD (0.5mg.ml) or 50μl Ndc10 CTD (0.5mg/ml). As controls the same

amount of core or different Ndc10 constructs with added reaction buffer were used. Everything else was conducted as described above.

2.8.2 ITC

Isothermal titration calorimetry (ITC) is a quantitative method to study protein-protein interactions. It utilises thermal changes, which occur during a binding reaction - heat is either produced (exothermic) or consumed (endothermic). Both proteins, Scm3 and Ndc10 CTD/CTDshort were purified as described in 2.5.1. The last purification step, the size exclusion chromatography, was conducted using the exact same buffer for all proteins. Scm3 peptides (Table 2) were synthesised by the Peptide Chemistry STP and dissolved also in the same buffer. Matching of the buffer is important as even slight differences in buffer composition lead to strong background signals. Ndc10 CTD or CTDshort was pipetted into the experiment cell at a concentration of 25μM and Scm3 or Scm3 peptides were prepared in the syringe at 221μM. Measurements were done on a MicroCal iTC200 (Malvern). Both reference and experiment cell were heated to 16°C and 20 injections were conducted per run. As a control, Scm3 was titrated into buffer only. Data was analysed with Origin software.

Table 2: Scm3 peptides

N-terminal peptides used for ITC measurements. Mutated residues are highlighted in bold. Hydrophobicity and pI were calculated using ProtParam webserver. The former is a grand average of hydrophobicity values, calculated by adding hydropathy values of all residues and dividing by the total number of residues (therefore, the more positive this score the more hydrophobic).

Name	Sequence	Hydrophobicity	pI
WT	KTNKKISKRRSLKNLHGALKGLLKES	-0.989	11.3
Mutant 1	KTNKKISKRRSLKNLHGA**A**KGLLKES	-1.063	11.3
Mutant 2	KTNKKISKRRS**AA**NLHGALKGLLKES	-0.852	11.26
Mutant 3	KTNKKISKRRSLKNLHGALKG**AA**KES	-1.137	11.3
Mutant 4	KTNKKISKRRS**AA**NLHGA**A**KG**AA**KES	-1.074	11.26

2.8.3 Crosslinking MS

Crosslinking coupled with mass spectrometry is an increasingly used method, to get a better understanding of protein/protein interactions, especially if some other

structural or biochemical data is present and can be used for validation. Both CBF3 and core complex were purified as described in 2.5.2. Straight after size-exclusion chromatography protein at 0.38mg/ml was crosslinked with DSSO, a crosslinker which can be cleaved in a mass spectrometer and gives a characteristic pattern used to identify crosslinked peptides. Therefore, 80µl of protein was mixed with 0.8µl of DSSO (Thermo Fisher Scientific; one vial is dissolved in 25.7µl of DMSO) and incubated at 37°C, 700rpm for 30min. The mixture is subsequently quenched by addition of 0.25% final concentration of hydroxylamine for 15min, 37°C, 700rpm and stored at -20°C. All further sample preparation, mass spectrometry and data processing were done solely by the Proteomics STP.

2.8.4 Peptide array

A peptide array can be used to identify potential stretches of residues important for protein/protein interaction. The array consists of a collection of peptides from protein A, whereas thousands of copies of each peptide are immobilised on a membrane in spots next to each other. This membrane is then incubated with protein B, to facilitate the binding to spots of peptides important for the interaction. After excess of protein B is washed off, it is detected with an appropriate antibody.

A peptide array membrane of Ndc10 was ordered from the Francis Crick Institute's Peptide Chemistry STP. It was designed to contain spots of 20 amino acid-long peptides, which are shifted by one amino acid each (e.g. spot 1: residue 1 to 20; spot 2: residue 2 to 21, etc.). Due to the size of Ndc10, two membranes were needed to accommodate all the peptides and two copies of each membrane were produced to allow for a control. Before use the membranes were prepared by incubation with stripping buffer B for 1hr at RT on a shaking platform, followed by washing with H_2O for three times. Membranes are then equilibrated with incubation buffer, before adding 200µg of Strep-tagged core complex per membrane and incubation overnight at 4°C on a shaking platform. Control membranes are incubated with buffer only. To facilitate detection of bound core complex, membranes are washed three times with incubation buffer plus 0.1% Tween-20 and blocked for 30min at RT with incubation buffer plus 0.1% Tween-20 and 2% milk. A 1:7000 dilution of primary strep-tag antibody (Bioworld Technologies) in incubation buffer plus 0.1% Tween-20 and 0.5% milk is added and incubated for 1hr at RT shaking. Unbound primary antibody is

washed off for three times with incubation buffer plus 0.1% Tween-20, before a 1:7000 dilution of secondary Anti-mouse antibody (GE Healthcare) in incubation buffer plus 0.1% Tween-20 and 0.5% milk is added and incubated for 1hr at RT, shaking. The membranes are subsequently washed for three times, before detection. Membranes are placed on clear film in a western blotting cassette, whereas care is taken to not let them dry. Per membrane, 5ml of freshly mixed Amersham ECL™ Western Blotting Detection reagents (GE Healthcare) solution is added. In the dark room, a photosensitive film is placed on the membranes and exposed for 30sec, before being developed in an automated developer. For reuse of the membranes, they can subsequently be stripped and stored dry in -20°C. Therefore, membranes are once washed with H_2O and subsequently incubated with stripping buffer A for 1hr at 55°C, 100rpm before being washed with H_2O again. Membranes are then incubated in stripping buffer B at room-temperature for 1hr on a shaking platform, washed once with H_2O, once with 20% EtOH and are then air-dried. Dry membranes can be stored at -20°C wrapped in filter paper and plastic.

Stripping buffer A:	**Stripping buffer B:**	**Incubation buffer:**
8M Urea	50% EtOH	10mM HEPES, pH 7.5
1% SDS	10% glacial acid	300mM NaCl
100mM β-mercaptoethanol		1mM DTT

2.8.5 EMSA

Electrophoretic mobility shift assays (EMSAs) are used to study Protein/DNA interactions. DNA and protein are mixed and run on an agarose or polyacrylamide native gel. Due to the increased size of the protein/DNA complex one can detect a shift of the DNA if compared to free DNA. In this thesis, solely agarose gels were used, as both CBF3 and core complex were unable to penetrate polyacrylamide gels, possibly due to their size. 10µl of samples containing protein and DNA, being either λ-Protein Phosphatase (NEB) treated or untreated, were mixed with 0.5µl of 5x DNA loading buffer blue (Bioline) and immediately loaded onto a precooled 0.8% agarose gel (without added EtBr). The gel was placed on ice and run at 80V for 20min in chilled 1x TBE buffer. DNA is stained with EtBr solution (50ml H_2O plus 2µl EtBr) for

20-30min and imaged with UV light with a Gel Doc™ EZ Imager (BioRad). Subsequently, it can be stained with InstantBlue (Expedeon) to visualise proteins.

Phosphatase reaction:

Core/CBF3 (0.3mg/ml)	5µl
Reaction buffer	2.5µl
100mM $MnCl_2$	1µl
λ phosphatase	0.5µl
56bp CEN3 (0.01 µg/µl)	1µl
total	10µl

Reaction buffer:

10mM HEPES, pH 7.5
300mM NaCl
1mM DTT

2.8.6 Analytical size exclusion chromatography

Size exclusion chromatography separates proteins according to their size, as well as shape and can therefore be used to determine if two proteins of interest are interacting. If there is an interaction, the size will increase significantly, given the two interaction partners are big enough. This will result in a shift of elution towards bigger size/earlier elution time.

Scm3 and Ndc10 CTD: 337µg Scm3 (full length) and 256.5µg of Ndc10 CTD were mixed in a total volume of 0.5ml. Protein mixture was injected and run on a Superdex 200, 10/300 GL (GE Healthcare) column, due to their expected size of ~120kDa (for a Ndc10 CTD homodimer plus a Scm3 monomer) or ~146kDa (for a Ndc10 CTD homodimer plus a Scm3 homodimer).

Scm3 and Ndc10 CTDshort: 750µg Scm3 (full-length) and 810µg CTDshort were mixed in a total volume of 0.5ml. As the expected size of a Scm3 or a CTDshort homodimer is ~50kDa, a Superdex 75, 10/300 GL (GE Healthcare) column was used.

2.9 Structural studies

As explained in 1.6, different methods exist to elucidate a structure of a protein. Here, both crystallisation trials, as well as single particle cryoEM was used.

2.9.1 Crystallization trials

Crystallisation of a protein or protein complex requires a pure and concentrated sample and a condition of precipitants to force crystallisation and prevent aggregation. The condition needed is not known and therefore, extensive trials are conducted with commercial or home-made screens. Once an initial hit is found, these conditions can be optimised until a crystal is achieved which gives a sufficient diffraction pattern. Here, three different proteins/complexes were tested for crystallisation: the CBF3 core (wild-type or Gal4Δ) complex, Ndc10 CTDshort and Ndc10 CTD bound to Scm3 NTD. Concentrated proteins were set up with different screens (see Table 3, Table 4, Table 5 and Table 6 for details) using a Matrix Hydra (Thermo Scientific) to suspend reservoir solutions and a Mosquito HTS (TTP Labtech) to pipette 400nl drops of protein and reservoir. Screens were kept and imaged in a Rock Imager (Formulatrix) either at room temperature or 4°C and drops were inspected using the Rockmaker software (Formulatrix). Optimisation screens were designed with the same and mixed using a Formulator (Formulatrix). Crystals were cryo-protected with a range of protectants, frozen and stored in LN_2. Diffraction pattern was recorded at Diamond Light Source.

Table 3: Screens used for crystallisation trials of CBF3 core

Protein concentration	Screen	Drop size	Ratio protein:reservoir	Incubation temperature
1.5 mg/ml *	Natrix	400nl	1:1	RT and 4°C
1.5 mg/ml *	Proplex	400nl	1:1	RT and 4°C
1.2 mg/ml	PACT	400nl	3:1	RT
1.2 mg/ml	JSCG+	400nl	3:1	RT
1.2 mg/ml	Morpheus	400nl	3:1	RT
1.2 mg/ml	PGA	400nl	3:1	RT
1.2 mg/ml	Proplex	400nl	3:1	RT
1.65 mg/ml *	PACT	400nl	3:1	RT
1.65 mg/ml *	JCSG+	400nl	3:1	RT
1.65 mg/ml *	Morpheus	400nl	3:1	RT
1.65 mg/ml *	$(NH_4)_2SO_4$	400nl	3:1	RT

* marks occasions which were set up without and with 27bp CEN3 DNA in a 1:1 molar ratio.

Table 4: Screens used for crystallisation trials of CBF3 core Gal4Δ

Protein concentration	Screen	Drop size	Ratio protein:reservoir	Incubation temperature
2 mg/ml	Natrix	400nl	1:1 and 3:1	RT
2 mg/ml	Proplex	400nl	1:1 and 3:1	RT
2 mg/ml	Clear Strategy I	400nl	1:1 and 3:1	RT
2 mg/ml	JSCG+	400nl	1:1 and 3:1	RT

Table 5: Screens used for crystallisation trials of Ndc10 CTDshort

Protein concentration	Screen	Drop size	Ratio protein:reservoir	Incubation temperature
10 and 20 mg/ml	JSCG I	400nl	1:1	RT and 4°C
10 and 20 mg/ml	JSCG II	400nl	1:1	RT and 4°C
10 and 20 mg/ml	JSCG III	400nl	1:1	RT and 4°C
10 and 20 mg/ml	JSCG IV	400nl	1:1	RT and 4°C
10 and 20 mg/ml	Natrix	400nl	1:1	RT and 4°C
10 and 20 mg/ml	PACT	400nl	1:1	RT and 4°C
10 and 20 mg/ml	PEG I	400nl	1:1	RT and 4°C
10 and 20 mg/ml	PEG II	400nl	1:1	RT and 4°C
10 and 20 mg/ml	Wizard I-II	400nl	1:1	RT and 4°C
10 and 20 mg/ml	Wizard III-IV	400nl	1:1	RT and 4°C

Table 6: Screens used for crystallisation trials of Ndc10 CTD and Scm3 NTD

Protein concentration	Screen	Drop size	Ratio protein:reservoir	Incubation temperature
5 and 2.5 mg/ml	JSCG I	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	JSCG II	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	JSCG III	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	JSCG IV	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	Natrix	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	PACT	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	PEG I	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	PEG II	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	Proplex	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	Wizard I-II	400nl	1:1	RT and 4°C
5 and 2.5 mg/ml	Wizard III-VI	400nl	1:1	RT and 4°C
6.8 mg/ml	JSCG+	400nl	1:1 and 3:1	RT and 4°C
6.8 mg/ml	PACT	400nl	1:1 and 3:1	RT and 4°C
6.8 mg/ml	Pi-PEG	400nl	1:1 and 3:1	RT and 4°C
6.8 mg/ml	MIDAS	400nl	1:1 and 3:1	RT and 4°C
6.8 mg/ml	Morpheus	400nl	1:1 and 3:1	RT and 4°C
6.8 mg/ml	PGA	400nl	1:1 and 3:1	RT and 4°C
6.8 mg/ml	PEG I	400nl	1:1 and 3:1	RT and 4°C
6.8 mg/ml	PEG II	400nl	1:1 and 3:1	RT and 4°C

2.9.2 Negative stain EM

Straight after size exclusion chromatography CBF3 and core complex was, if necessary, diluted to a concentration of ~0.02 or 0.01mg/ml respectively. Copper 400 mesh coated with carbon (EM Resolution) were glow discharged with a K100X system (Quorum Technologies) for 30sec at 45mPa. 4μl of protein sample was pipetted on the carbon side of the grid and incubated for 1min. For staining four 60μl drops of 2% uranyl acetate are prepared on parafilm, nearly all of the protein solution is blotted away with filter paper (Whatman 2) and immediately after the grid is washed in one drop after the other for 10sec each. The grid is removed from the last drop and blotted dry with filter paper. After completely air-drying, it can be stored in a grid box indefinitely. Care has to be taken when handling the grid at all times, as for their fragile nature.
Negative stained grids were imaged with a 120kV G2 Spirit Twin TEM (FEI) at room temperature with a single-tilt side entry holder and a CCD camera. Core data was recorded manually and with a defocus range from -1 to -2μm.

2.9.3 CryoEM

Core and core SΔ was diluted to 0.1mg/ml straight after size exclusion. 56bp CEN3 DNA was added to the latter in a molar ratio of 1:1.2 (protein to DNA). The sample was frozen in liquid ethane using a Vitrobot or a manual plunge freezer. For the latter, protein was added to the glow discharged grids, held with a tweezer connected to the plunge freezer, incubated for 30sec and blotted with filter paper (Whatman 2) per hand. The Vitrobot was used as per instructions, at either 90% humidity or without humidifying at room temperature. A range of grids was tried for screening, namely C-flat (EMS), lacey carbon grids (EM Resolutions) and copper 1.2/1.3 400mesh or 2/2 200 mesh Quantifoil grids. After plunge freezing, grids were transferred and subsequently stored in LN_2.
Screening of ice thickness and particle appearance was conducted on a 120kV G2 Spirit Twin TEM (FEI) with a 626 Gatan side-entry cryo-holder in low-dose mode. In addition, screening was conducted at the Imperial College London's Electron Microscopy Center, with a CM200 TWIN (Phillips), a Tecnai G2 Spirit (FEI) and a 200kV Tecnai F20 G2 (FEI) all with a 626 Gatan side entry cryo-holder and a CCD camera. Once the right freezing conditions were established, data was collected on

Krios Titan microscopes (FEI) with a Gatan K2 camera both at eBIC (Diamond Light Source) and in-house. Data collection parameters for all three datasets are outlined in Table 7.

Table 7: CryoEM data collection parameters for all three datasets

	1st Dataset Core	**2nd Dataset Core**	**1st Dataset CoreS**
Microscope	Titan Krios (FEI)	Titan Krios (FEI)	Titan Krios (FEI)
Camera	K2 Summit (Gatan)	K2 Summit (Gatan)	K2 Summit (Gatan)
Voltage (keV)	300	300	300
Defocus range (μm)	-2 to -4	-1.5 to -3.5	-1.5 to -3.5
Magnification	130,000 x	130,000 x	130,000 x
Pixel size (Å/pixel)	1.08	1.078	1.08
Dose rate (e^-/pixel/s)	5.4	6.3	6.3
Movies	1608	8755	4000
Frames per movie	25	25	25
Total exposure time (sec)	10	10	10
Total accumulated dose (e^-/Å^2)	50	54	54
Particles after autopicking	~180k	~3.5 million	~890k
Particles contributing to final reconstruction	72,771	209,751	200,184
Sharpening B-factor (Å^2)	-150	-150	-150
Final resolution (Å)	5.7	3.6	3.9

2.9.4 Image processing

Negative stain data of the core complex was processed with EMAN2. Particles were picked automatically and quality control of picking was conducted manually. 2D classes were calculated and a subset of good classes was chosen. With an initial model derived from the lowpass-filtered Cep3Δ crystal structure, a 3D reconstruction was calculated.

The first core dataset was entirely processed within Relion2.0 using standard procedures (Scheres, 2016). In the case of both the 2nd core and core Skp1Δ dataset, movies were directly fed into Scipion1.1 and motion corrected with MotionCor2 on the flow by the cryoEM facility. The output was then fed into Relion2.1 and processed within Relion, as well as Cryosparc. 3D densities were manually inspected in UCFS Chimera (Pettersen et al., 2004). Local resolution calculation was done from within the Relion GUI with ResMap. Overall resolution was calculated by Fourier shell correlation (FSC) and the 0.143 criterion (Rosenthal and Henderson, 2003) and map

anisotropy was estimated using 3DFSC (Tan et al., 2017). Fitting of crystal structures was done manually in Chimera, treating the entire crystal structures as rigid bodies. Re-building of models was carried out in Coot (Emsley and Cowtan, 2004) and real-space refinement was done using phenix.refine (Adams et al., 2004).

2.10 *In vivo* analysis of Skp1

While *in vitro* studies of proteins are crucial to understand its functions, it does not always reflect on processes happening *in vivo*. For example, a protein might get phosphorylated by a specific kinase *in vitro*, however, in the cell they are spatially or temporarily separated and therefore this reaction never takes place. Therefore, a finding *in vitro* should always be confirmed or brought into context *in vivo*. Similarly, an observation *in vivo* could be misleading and one needs to confirm said observation *in vitro*. An example for the latter would be, a protein deleted *in vivo* showing differences in phosphorylation status of the cell, however, only *in vitro* studies with purified components can distinct between said protein being a kinase or phosphatase itself or having an indirect influence.

2.10.1 Tagging of endogenous proteins

Both constructed strains were of the W303a (MATa; ade2-1; his3-11,15; leu2-3,112; trp1-1; can 1-100; ura3-1) background and C-terminal tagging was performed by gene targeting using PCR products. Vectors containing either an PK6- or HA-tag, as well as a selective marker, were kindly provided from Frank Uhlmann's laboratory. The tag and selective marker are amplified with PCR using primers including overhangs complimentary to the very end, as well as just downstream of the gene of interest. The stop codon is excluded. PCR product is then transformed (see 2.3.1) and will integrate after the gene and therefore fuse the tag to the gene. Correct insertion is tested by integration PCR (see 2.3.2) with primers complimentary to a genomic sequence before and after the integration site, paired with primers complimentary to the inserted sequence. Unlike vectors for recombinant expression of genes of interest, those used for endogenous tagging have selective markers derived from a different species (e.g. *K.lactis*). This avoids targeting of the PCR product to the selective marker gene in the genome. Gene replacement to construct

the Skp1Δ strain was done using a modified version of the HA-tag vector, in which the Skp1 sequence, with the 37-64 loop deleted, was cloned fused to the HA-tag. Primers to amplify the product used for transformation were then designed to include the Skp1Δ gene and replace the entire endogenous gene. Insertion was tested by integration PCR, as mentioned above.

2.10.2 Cell cycle arrest and time-course experiment

To be able to test behaviour of a protein during the cell cycle one must synchronise all cells in a culture. Different protocols are available to arrest cells in different stages of the cell cycle. Here, cells were arrested in G1 by addition of alpha factor. After removal of this pheromone cells will synchronously move through one or two cell cycles and samples can be taken at specified time-points depending on the doubling time of the constructed strains. The doubling time was determined by growing an ONC in YP-media complemented with 2% glucose at 25°C and 200rpm, inoculating a 200ml day-culture to an OD_{600} of 0.2 and measuring cell density every hour after the culture has reached an OD_{600} of 0.3. The resultant growing curve is used to determine the doubling time in the log phase of growth.

A 450ml culture is inoculated in the evening and grown over night at 25°C, 200rpm in YP-media with 2% glucose to reach an OD_{600} of 0.2 to 0.3 the next morning. Alpha factor is added to an end concentration of 2µg/ml three times spaced by one hour each. 30mins after the third alpha factor dose, cells are visualised under the microscope to check for arrest. Arrested cells exhibit a distinctive shape, called shmoo. Once more than 90% of cells are shmooing, cells are washed with three times 400ml of fresh YP-media to remove any excess alpha factor. To do so cell suspension is poured in a sterile vacuum filter unit (Millipore) and whilst the media passes through the 0.22µm filter, cells are collected on top. Once all media has passed through wash media is added and also filtered through. After the washing, cells collected on the filter are resuspended in pre-warmed 450ml YP-media plus 2% glucose and transferred into a 2-litre conical flask. For three hours two samples, one for the CoIP and one for FACS analysis, are collected every 15min. After one hour cells are visualised under the microscope again, to check if cells have started to bud. Once more than 90% of cells are budding, another dose of alpha factor is added to the remaining cultures to arrest them in G1 again, after completing one cell cycle.

Another dose of alpha factor is added after 30min and cells are again tested for shmoo. In total 12 samples are collected.

2.10.3 Fluorescence-activated cell sorting (FACS)

To determine DNA content and therefore cell cycle stages of each time point, 1ml of cells are collected in a pre-chilled Eppendorf tube. Cells are pelleted for 1min, 4000g at 4°C. The media is decanted and 1ml of ice-cold 70% EtOH is added to fixate the cells. Samples are kept at 4°C for over the weekend until further processing. Cells are quickly spun down again and the supernatant is decanted. Cells are resuspended in 1ml of 50mM Tris/HCl, pH7.5 containing 0.1mg/ml RNase A (Sigma-Aldrich). Cells are incubated at 37°C, 200rpm overnight to digest all RNA. The next morning cells are pelleted and supernatant is decanted. Cells are resuspended in 400µl of FACS solution supplemented with 50µg/ml propidium iodide (1mg/ml solution in water, Sigma-Aldrich). Cell suspension is briefly sonicated with a Branson Sonifier 450 at step 6 for 5sec to prevent clumping of cells. Samples are then measured with a FACSCalibur (Becton Dickinson) according to standard protocol for haploid yeast cells. Data is visualised with FlowJo software.

FACS solution:

10ml Tris/HCl, pH 7.5 from 1M stock
2.1ml NaCl from 5M stock
3.9ml $MgCl_2$ from 1M stock
34ml H_2O

2.10.4 CoIP

At each time-point 30ml of cells are collected in a pre-chilled 50ml Sarsted tube and pelleted by centrifugation at 4000g, 5min at 4°C. Media is decanted and cells are resuspended in 1ml pre-chilled H_2O and transferred into 2ml cell breaker tubes. Cells are pelleted again via centrifugation, supernatant is decanted and cells are shock-frozen in LN2. After FACS measurement, showing a successful synchronisation/cell cycle passage, samples are processed for the CoIP.
Cells are resuspended in 400µl of lysis buffer with added benzonase and complete protease inhibitor, glass beads (Sigma) are added and cells are broken using a cell

breaker at 4°C. Cell breakage is determined with an optical light microscope and after more than 60% of cells are successfully broken lysate is collected and insoluble material as well as cell debris is separated by centrifugation at 13 000g, 10min at 4°C. Soluble supernatant is transferred into a fresh Eppendorf and protein concentration is measured with Bradford reagent and absorbance at 595nm. A total of 3.1mg of cell extract is used for the CoIP and the appropriate amount is transferred into a 1.5ml Costar tube (Corning) and the volume is normalised. A total cell extract sample is taken and mixed with SDS-loading buffer and frozen at -80°C. To preclear the cell extract for CoIP 20µl of unspecific antibody, coupled to Dynabeads is added and incubated on a shaker wheel at 4°C for 30min. Beads are collected on a magnetic rack and cleaned cell extract is transferred into a fresh Costar tube, mixed with 20µl of protein A dynabeads with coupled antibody and incubated on a wheel for 1hr, 4°C. Beads (bound fraction) are collected on a magnetic rack, washed three times for 10min, 4°C on the wheel and resuspended in 25µl of SDS-protein loading buffer and samples are frozen at -80°C.

Lysis buffer:
50mM HEPES, pH7.5
200mM NaCl
1mM DTT
50mM NaF
1mM $MgCl_2$
0.1mM Vanadate
10mM ß-glycerophosphate

2.10.5 Phos-tag SDS-PAGE and western blot

Frozen samples are thawed and briefly spun down. Total cell extract samples are boiled for 5min, 98°C, while bound fraction samples are only heated to 60°C for 5min to avoid excessive background signal from the heavy and light chain of the antibody. 10µl of samples are loaded onto a 12% Phos-Tag gel (see 2.7.5 for recipe) and run for 45min, 220V. Proteins are blotted on an activated PVDF membrane with the Mini Trans-Blot® Cell system (BioRad) using 1x transfer buffer (BioRad) chilled with a cooling element to avoid overheating. HA-tagged Skp1 is visualised via western blot (for procedure see 2.7.6) with a primary HA-tag antibody (Abcam) and secondary

anti-rabbit HRP conjugated antibody (GE Healthcare). Blots are developed in the dark room using an automated developer (Abnova) for 10sec each.

2.11 Protein constructs and DNA fragments used in this study

The following two tables summarise all DNA fragments used and all protein constructs described.

Table 8: Protein constructs used in this study

All tags contained a TEV site to be cleavable with TEV protease, except for sumo-tags, which can be cleaved with sumo protease without any additional introduced residues.

Name	Mutation/Deletion	tag
CBF3SΔ	full complex, including Skp1 residues 37 to 64 deleted	C-terminal 2xstrep on Cep3
Cep3	n.a.	C-terminal 2xstrep or HA and C-terminal 2xstrep
Cep3 2xmut or M3	combination of Cep3 LL_SA and Y_S	HA and C-terminal 2xstrep
Cep3 LL_SA or M1	L319 and L320 mutated to serine and alanine, respectively	HA and C-terminal 2xstrep
Cep3 SA	S575 mutated to alanine	C-terminal 2xstrep on Cep3
Cep3 SE	S575 mutated to glutamic acid	C-terminal 2xstrep on Cep3
Cep3 Y_S or M2	Y325 mutated to serine	HA and C-terminal 2xstrep
core	core complex (Cep3, Ctf13, Skp1)	C-terminal 2xstrep on Cep3
core Gal4Δ	core complex with Cep3 residues 1-46 deleted	C-terminal 2xstrep on Cep3
coreSΔ	core complex, including Skp1 residues 37 to 64 deleted	C-terminal 2xstrep on Cep3
Ctf13	n.a.	n.a.
Mif2-530	residues 1 to 255, and 531 to end deleted	N-terminal his
Mif2-end	residues 1 to 255 deleted	N-terminal his
Ndc10	na	na or N-terminal HA, or C- or N-terminal 2xStrep
Ndc10 centralCTD	residues 551-745	N-terminal his tag
Ndc10 CTD	residues 551-end	N-terminal sumo tag or n.a.
Ndc10 CTDshort	residues 746-end	N-terminal his tag
Ndc10 degronMut	degron mutated, as in	N-terminal his
Ndc10 NTD	residues 1-551 (this construct was available already purified)	n.a.
Ndc10 Δ10	last 10 amino acids deleted	N-terminal his
Scm3	n.a.	N-terminal his tag
Scm3 NTD	residues 1 to 95	N-terminal his tag
Skp1	n.a.	n.a. or N-terminal 2xstrep or HA
Skp1 NY_KK	N139 and Y140 mutated to two lysines	N-terminal 2xstrep or HA
Skp1 SA	Phosphomutant: S162 mutated to alanine	N-terminal 2xstrep
Skp1 SATA	Phosphomutant: S162 and T177 both mutated to alanine	N-terminal 2xstrep

Skp1 SE	Phosphomutant: S162 mutated to glutamic acid	N-terminal 2xstrep
Skp1 SETD	Phosphomutant: S162 and T177 mutated to glutamic and aspartic acid, respectively	N-terminal 2xstrep
Skp1 TA	Phosphomutant: T177 mutated to alanine	N-terminal 2xstrep
Skp1 TD	Phosphomutant: T177 mutated to aspartic acid	N-terminal 2xstrep
SKP1Δ	residues 37 to 64 deleted	N-terminal 2xstrep

Table 9: Sequences of DNA fragments used in this study

Highlighted features are the conserved CCG triplet (red), the second suggested TGT triplet (blue) and CDEI and CDEIII (bold). All fragments were used as double stranded DNA. In the case of 56bp and 33bp fragments, both strands were purchased from Sigma, mixed in an equal molar ratio, heated to 98°C and cooled stepwise to facilitate annealing. Longer DNA pieces were made as described in 2.6.1.

Name	Sequence
56bp WT CEN3	5'-TATTAG**TGTATTTGATTTCCGAAAGTTAAAA**AAGAAATAGTAAGAAATATATATTT-3'
33bp wild-type CEN3	5'-AAATATTAG**TGTATTTGATTTCCGAAAGTTAAA**-3'
33bp CCG mutant CEN3	5'-AAATATTAG**TGTATTTGATTT*GAT*AAAGTTAAA**-3'
33bp TGT mutant CEN3	5'-AAATATTAG***AAC*ATTTGATTTCCGAAAGTTAAA**-3'
33bp double mutant CEN3	5'-AAATATTAG***AAC*ATTTGATTTGATAAAGTTAAA**-3'
147bp Widom 601 DNA	5'-ATCGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCGAT-3'
147bp CEN3 DNA	5'-AAATAGTACAAATAA**GTCACATG**ATGATATTTGATTTTATTATATTTTTAAAAAAAGTAAAAAATAAAAAGTAGTTTATTTTTAAAAAATAAAATTTAAAATATTAG**TGTATTTGATTTCCGAAAGTTAAAA**AAGAAATAGTAAGAA-3'
185bp CEN3 DNA	5'- CAAATAA**GTCACATG**ATGATATTTGATTTTATTATATTTTTAAAAAAGTAAAAAATAAAAAGTAGTTTATTTTTAAAAAATAAAATTTAAAATATTAG**TGTATTTGATTTCCGAAAGTTAAAA**AAGAAATAGTAAGAAATATATATTTCATTGAATGGATATATGAAACGTTTACTGGTGGAAG -3'

2.12 Bioinformatics

Multiple sequence alignments were done using the webserver Clustal Omega (Sievers et al., 2011; https://www.ebi.ac.uk/Tools/msa/clustalo/), as well as NCBI web tools.

Graphical rendering of sequence conservation and secondary structure elements derived from a pdb file were created with the ESPript webserver (Robert and Gouet, 2014; http://endscript.ibcp.fr/ESPript/ESPript/index.php).

2.13 Solutions and buffers

LB media (or agar)

10g/l bacto Tryptone, 5g/l yeast extract, 10g/l NaCl, (15g/l difco agar)

PBS

8g/l NaCl, 0.25g/l KCl, 1.437g/l Na_2HPO_4, 0.25g/l KH_2PO_4

SDS/Protein loading buffer

10% SDS, 10mM β-mercaptoethanol, 20% Glycerol, 0.2M Tris-HCl pH 6.8, 0.05% Bromophenol blue

SDS-running buffer

25mM Tris HCl, 200mM glycine, 0.1% SDS

SOC media

20g/l bacto tryptone, 5g/l yeast extract, 0.584g/l NaCl, 0.186g/l KCl, 2.033g/l $MgCl_2.6H_2O$, 2.464g/l $MgSO_4.7H_20$, 2.603g/l glucose

1xTBE

10.8g/l Trizma base, 5.5g/l boric acid, 2ml 0.5M EDTA

YN-media (or agar)

67g/l Difco yeast nitrogen base (w/o) amino acids), (20g/l difco agar)

YP-media (or agar)

10g/l yeast extract, 20g/l bacto peptone, (20g/l Difco agar)

Chapter 3. Results 1: Establishment of a successful expression system for CBF3

Despite its crucial role in budding yeast kinetochore establishment and its requirement for *in vitro* kinetochore reconstitution, it has remained difficult to express and purify CBF3, especially in high yields needed for structural studies. These difficulties most likely arise from posttranslational modifications essential for complex assembly, as well as from the necessity of assembly factors (see introduction 1.4.1). The first and very crucial step in my work, therefore, was to establish a reliable and high-yield expression system and a working purification pipeline for CBF3. Although Cep3, Skp1, and subdomains of Ndc10 can be readily expressed in *E.coli*, it has been proven difficult to express Ctf13, or any sub complex containing Ctf13 in bacteria. To circumvent this problem, I tested co-expression of all four subunits both in budding yeast itself, as well as with the baculovirus/insect cell system.

3.1 Co-expression in insect cells

The baculovirus/insect cell expression system is often chosen for recombinant expression of protein complexes and when posttranslational modifications, such as phosphorylation, are needed. It also offers a high level of protein expression, which makes it a valuable tool for structural studies.

To test if the complex can be expressed and also formed in insect cells, codon-optimised genes were cloned in the Multibac system vectors (see 2.4.1), the virus was produced and a 4-litre culture infected. However, no expression of CBF3 and only a weak expression of the tagged Cep3 subunit could be detected (Figure 11A). This points towards the necessity of yeast-specific assembly factors needed for complex formation, with Sgt1 being a likely possibility (see introduction 1.4.1).

Simultaneous expression of Sgt1, however, was not tested, as below described co-expression in budding yeast was successful.

3.2 Co-expression in budding yeast

To circumvent issues with either lack of assembly factors or essential posttranslational modifications, a co-expression in budding yeast itself was

attempted. Therefore codon-optimised genes were cloned into inducible yeast expression vectors kindly provided by John Diffley. All four proteins could be detected with SDS-PAGE/Coomassie stain after affinity purification, however, still with a rather low yield (Figure 11B).

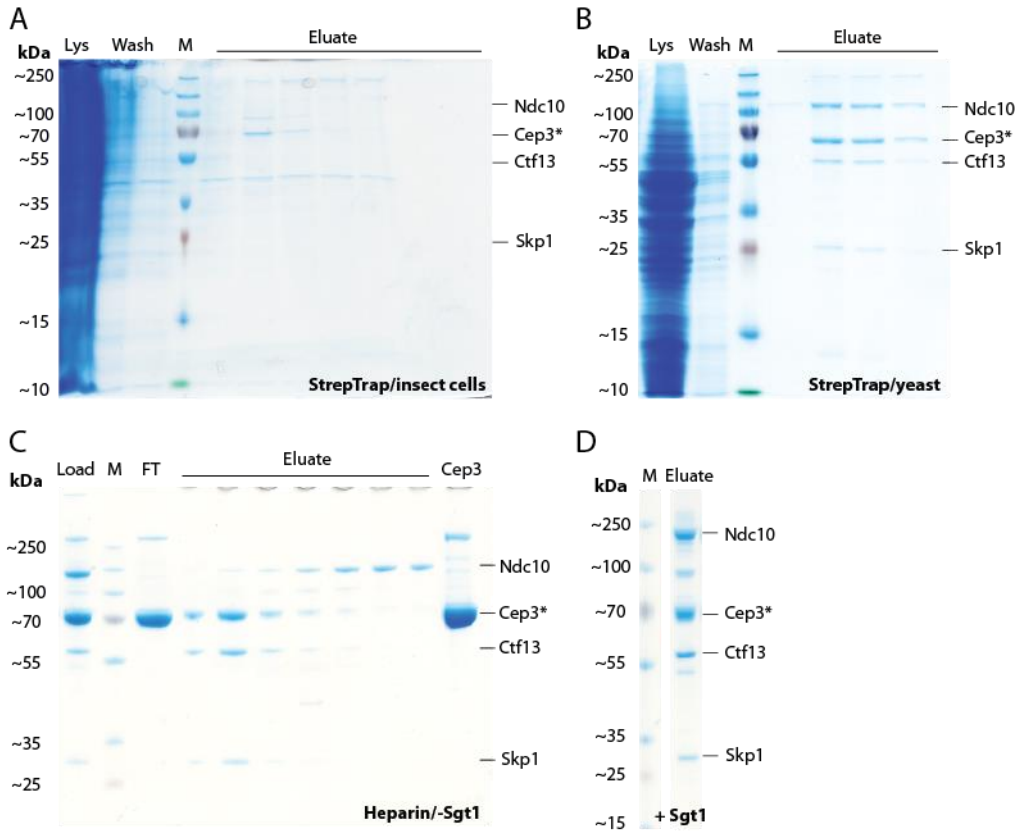

Figure 11: CBF3 expression tests

SDS-PAGE/Coomassie stain of a StrepTrap purification from CBF3 expressed in insect cells (**A**) and yeast (**B**). In both cases the full lysate (Lys), wash fraction (Wash) and the eluate fractions (Eluate) were run, as well as a molecular weight marker (M). * indicates that Cep3 is C-terminal tagged. **C** SDS-PAGE/Coomassie stain of a typical Heparin purification with yeast expressed protein without simultaneous co-expression of Sgt1. Excess Cep3, found in the load, does not bind to Heparin column and is therefore in the flow-through (FT), whereas the core complex and Ndc10 both bind, but elute separately from each other. The last lane shows concentrated Cep3 from the flow-through, which was used for further purification of Cep3 alone. **D** SDS-PAGE/Coomassie stain of a StrepTrap elution of CBF3 co-expressed with Sgt1, showing no excess of Cep3 and a higher overall yield of the other CBF3 subunits.

As all genes are under the control of the Gal4 promoter, endogenous Gal4 can become a limiting factor of expression. Therefore, Gal4 was co-expressed

simultaneously, which indeed improved the yields considerably (Figure 11C). Similarly, as an excess of the tagged Cep3 subunit was observed, it was tested if co-expression of Sgt1, a known assembly factor (see introduction 1.4.1), could further improve the yield and remove any excess of the tagged subunit Cep3, as shown in Figure 11D. Notably, the addition of phosphatase inhibitor in the lysis buffer was essential to be able to pull out the full complex, as well as a to achieve a good yield.

3.3 Purification of CBF3

Although all four subunits were present in the initial affinity purification, Ndc10 seemed to be able to readily dissociate from the core complex, as it was also found in the wash, suggesting a weak interaction (Figure 12A).

Further purification was achieved by heparin, to exclude any excess of Cep3 as well as impurities (Figure 12B). Interestingly, this step also led to the separation of Ndc10 and the core complex, again speaking for a weak interaction between the two. However, this separation proved to be a useful way to separately purify the core complex and full-length Ndc10. The latter on its own is very unstable and readily degrades. Furthermore, any attempts to express and purify either N- or C-terminal tagged Ndc10 alone in budding yeast proved to be unsuccessful. As described in the introduction (see 1.4.1) Ndc10 features a degron motif at the very C-terminus. To test if the protein can be stabilised for purification by inactivation of this motif, both a deletion or a mutation of this motif, as described in Furth et al., 2011 and Alfassy et al., 2013, was cloned and tested. However, there was still no notable expression of full-length Ndc10 alone in budding yeast.

For further purification of the full complex, Ndc10 and the core is pooled together and subjected to anion exchange chromatography. For purification of the core complex only, Ndc10 is omitted at this step. Anion exchange chromatography is crucial to separate co-purified DNA, as well as some prominent protein impurities (Figure 12C). Finally, size exclusion chromatography is performed, which separates core and full CBF3 complex due to their size difference. The unstable nature of Ndc10 proved to be the limiting factor for the yield of the full complex, and the stable core complex is therefore present in excess. Both complexes, however, show a promising elution profile and expected stoichiometry of 2xNdc10, 2xCep3, 1xCtf13 and 1x Skp1, as judged by band intensities after SDS-PAGE/Coomassie stain (Figure 12D).

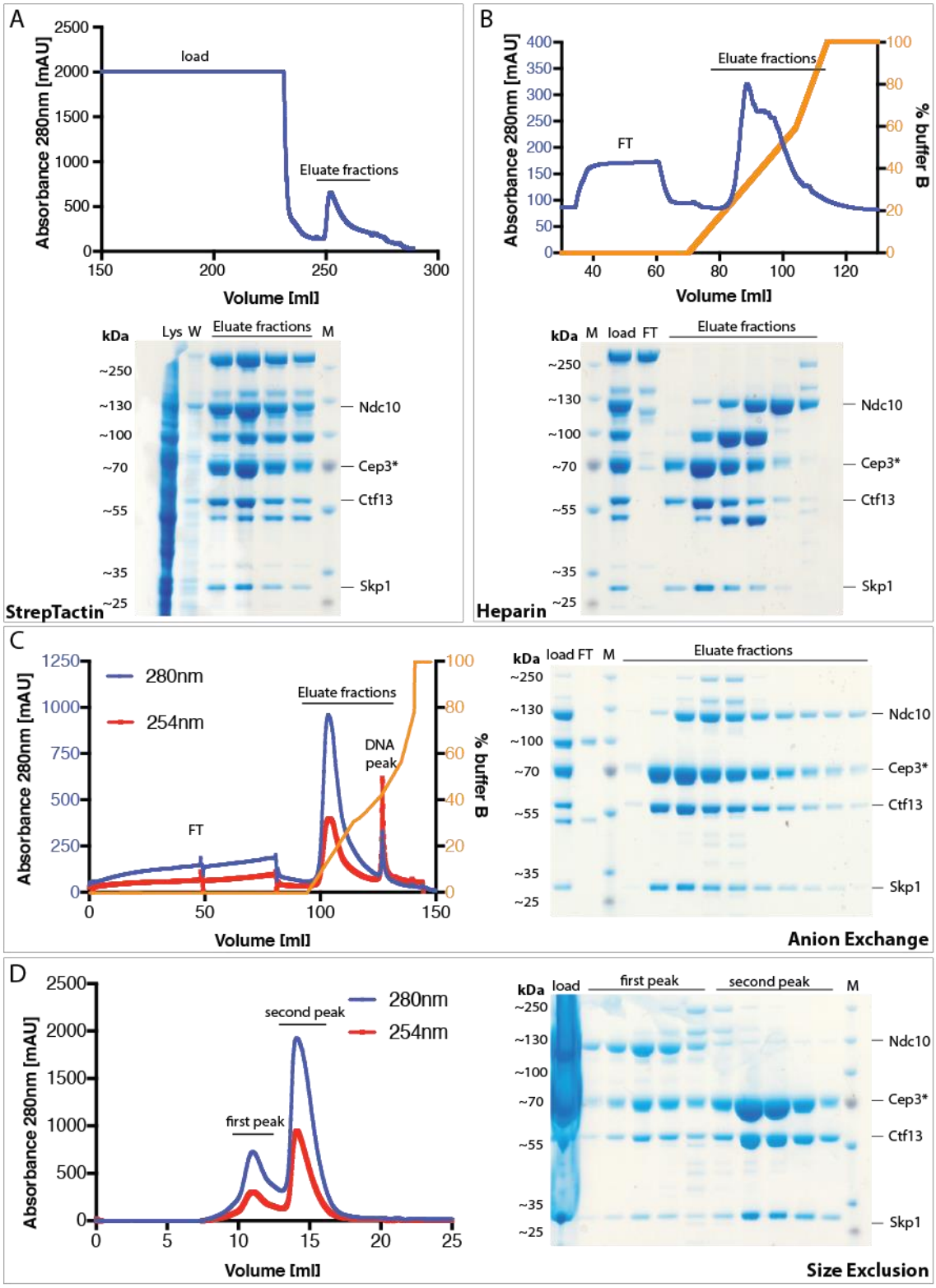

Figure 12: Purification of CBF3

A Typical elution profile of initial Strep-Tactin affinity purification and SDS-PAGE of lysate (Lys), wash (W) and elution fractions. The whole elution peak, as indicated in A was pooled for further purification. **B** Typical elution profile of heparin column and SDS-PAGE of load,

wash (W) and elution fractions. As in the first step, the whole elution peak as indicated in B was pooled. **C** Elution profile of anion exchange chromatography, as well as SDS-PAGE of load, flow-through (FT) and elution peak fractions. At this step the protein peak was pooled without including the DNA peak at ~130ml. **D** Final step; size exclusion chromatography with a Superose 6 10/300 column and SDS-PAGE of load and the two peak fractions corresponding to the full complex (first peak) and core complex (second peak). M… molecular weight marker

Both full and core complex exhibit sensitivity to salt concentration and start aggregating in buffers containing less then 250mM NaCl. The final buffer therefore was 10mM HEPES, pH 7.5, 300mM NaCl and 1mM DTT.

The stoichiometry was further verified by molecular weight measurement with SEC-MALS. The predicted masses of the full and the core complex, in above-mentioned stoichiometry are ~445kDa and ~220kDa respectively. SEC-MALS analysis of the core complex consistently showed an experimental mass of ~230kDa, supporting the predicted mass (Figure 13A). Analysis of the full complex produced somewhat less consistent results, however, it showed a mass between ~450 to ~500kDa (Figure 13B), also consistent with the predicted stoichiometry.

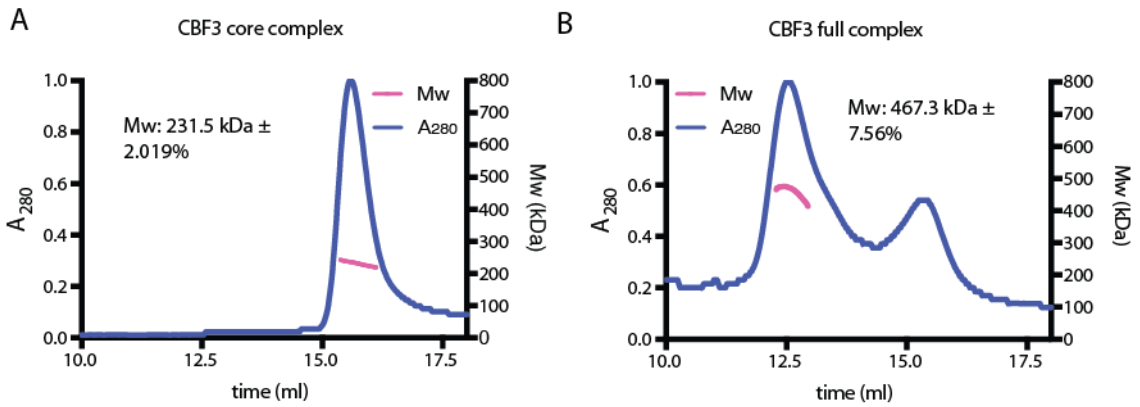

Figure 13: SEC-MALS of the core and full CBF3 complex

SEC-MALS curve of one of two measurements of the core complex (**A**) and the full CBF3 complex (**B**). Measurements were conducted in duplicates and whilst the core complex had a consistent reading, the full complex did not and the second measurement showed a mass of ~500kDa.

Chapter 4. Results 2: Structural studies of CBF3

After the successful establishment of an expression and purification system and given the promising size exclusion profiles of both the full and core CBF3 complex, structural elucidation of the CBF3 complex was attempted. Whereas in the past X-ray crystallography dominated as a method for structure determination, more recently cryoEM has emerged as a more and more powerful tool, especially if one is interested in bigger protein complexes with low yields. Given the, for cryoEM, rather small size of the core (~220kDa), both crystallisation trials, as well as negative stain EM screening was performed. On the other hand, due to limiting yield and the instability of Ndc10, only negative stain EM was considered for the full CBF3 complex.

4.1 Crystallisation trials of the core complex

Core complex was concentrated to 1.2-1.5mg/ml and a variety of standard crystallisation screens were set up with and without 27bp CEN3 dsDNA. Further concentration was not possible due to aggregation. Small crystalline needles could be observed after 24 hrs in a few conditions only without DNA and at RT (Figure 14A and B). Optimisation of crystal conditions was attempted, with limited success and SDS-PAGE analysis of the crystals showed that Cep3 and Ctf13 degrade almost completely and only Skp1 remains (Figure 14C-I). Given that the flexible Gal4-domains of Cep3 have to be removed in order to crystallise Cep3 alone (see introduction 1.4.1), the same construct was used, to co-express a CBF3 complex lacking these domains (henceforth called CBF3 Gal4Δ or core Gal4Δ). Co-expression and purification were identical to the full length CBF3. Furthermore, core Gal4Δ could be further concentrated, to about 2mg/ml, before aggregation was observed, indicating a more stable nature of this construct. Crystallisation trials, however, did not show any improvement and no crystals could be obtained. Trials were concluded, as the core complex appears to be too unstable for crystallisation, as well as tending to aggregate hindering protein concentration.

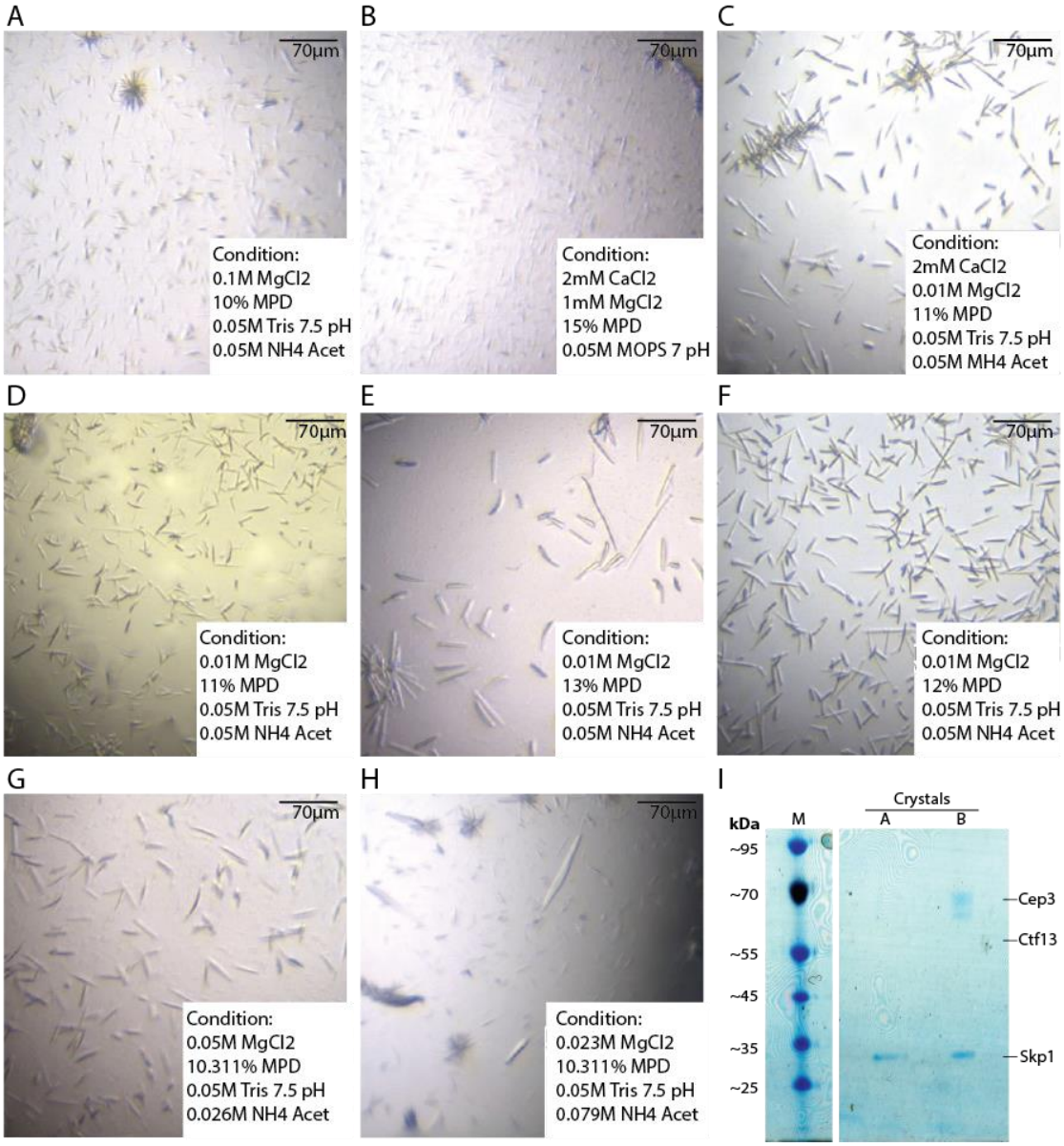

Figure 14: Core initial crystalline needles and optimisation

A-B Two initial hits of crystalline needles with 1.5mg/ml protein, without DNA and incubated at RT. Conditions are specified in the images. **C-H** Crystals after optimisation with conditions specified in the images. **I** SDS-PAGE/Coomassie stain of pooled crystals from conditions without (A) and with (B) precipitation. Usual migration height of the three proteins is indicated.

4.2 Negative stain EM of CBF3 and core complex

Given the unsuccessful crystallisation trials, both the full CBF3 and core complex were subjected to negative stain EM, which is a quick and easy method to test for suitability for structural elucidation by cryoEM. To be successful candidates,

proteins/protein complexes need to form distinct particles, with no or little orientation bias.

Dilutions of CBF3 and core were stained with uranyl acetate straight after size exclusion chromatography and imaged using a 120kV TEM. Whilst images of the core complex showed nice, uniform and distinct particles (Figure 15A), the whole complex looked less promising. Indeed, it seemed to be comprised of core particles and additional diffuse density, which must account for unstructured Ndc10 (Figure 15B). Buffer optimisation was attempted to improve CBF3 particles. Neither pH changes, nor salt concentration or salt type, nor type of buffer made any improvements, judged by both size-exclusion profile, as well as negative stain EM. This might not be very surprising given the predicted high disorder of Ndc10 especially in its C-terminal domain (Perriches, 2014). Attempts to improve full CBF3 particles by either changing Ndc10 constructs or binding it to centromeric DNA are described in 5.3.

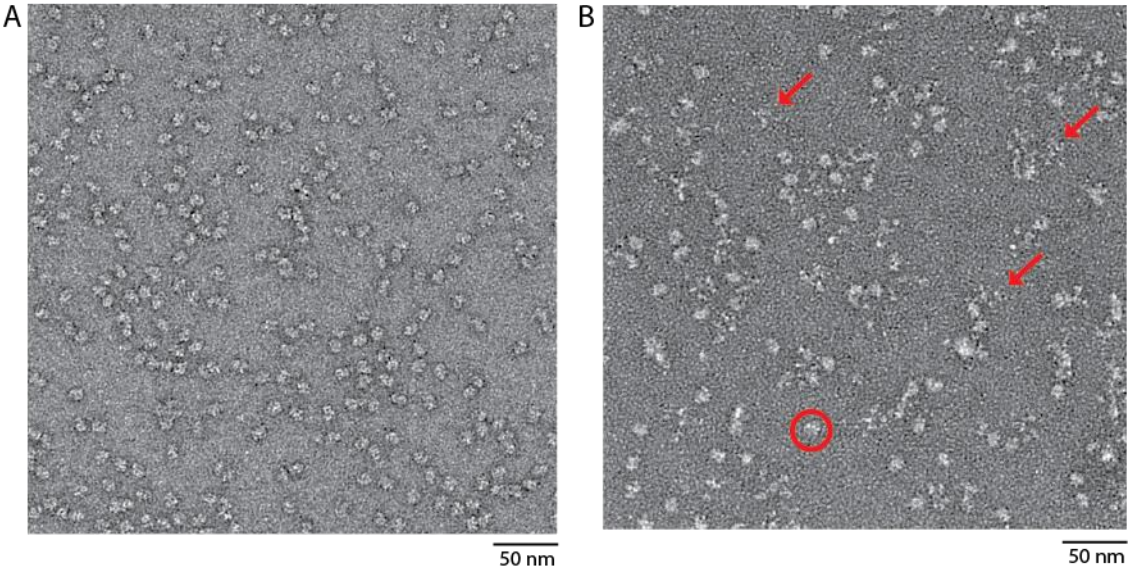

Figure 15: Negative stain analysis of CBF3 core and full complex

Typical micrograph of (**A**) core and (**B**) full CBF3 complex. The latter is comprised out of core particles (red circle) and additional diffuse density, assumable unstructured Ndc10 (red arrows).

To get a better understanding of the suitability of the core for high-resolution structural cryoEM, a small dataset of ~300 micrographs of negative stained core was collected with the in-house 120kV screening microscope. Particles were picked and

processed with EMAN2. 2D classes looked promising with many different orientations present, and a ~40Å 3D model could be generated by using the low-pass filtered Cep3 crystal structure (2VEQ) as an initial model (Figure 16A and B).

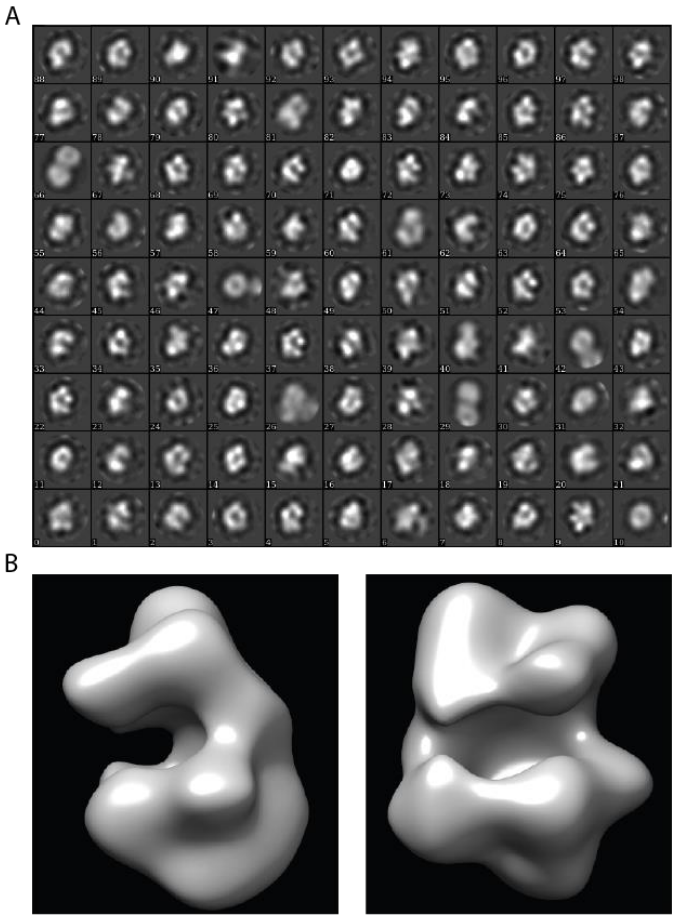

Figure 16: Image processing of negative stained core complex
A Unsorted 2D classes of core complex as processed with EMAN2. **B** 3D reconstruction of negative stained core complex in a side (left image) and front (right image) orientation at a resolution of about 40 Å, revealing the overall shape of the core complex.

4.3 CryoEM analysis of the core complex

After promising negative stain imaging and reconstruction of the core complex, freezing conditions had to be established. As described in introduction (see 1.6.2), one needs a protein/protein complex to be frozen in a very thin amorphous ice layer, without any major orientation bias, to be able to perform data collection as well as 3D reconstruction. Therefore, protein was prepared as described before and cryo grids were frozen with both a FEI Vitrobot, as well as a manual plunge freezer. Whilst grids prepared with the Vitrobot exhibited nice uniform and thin ice, the protein was either absent or could only be detected as dark precipitate (Figure 17A), pointing to a potential sensitivity of the core complex with the air-water interface.

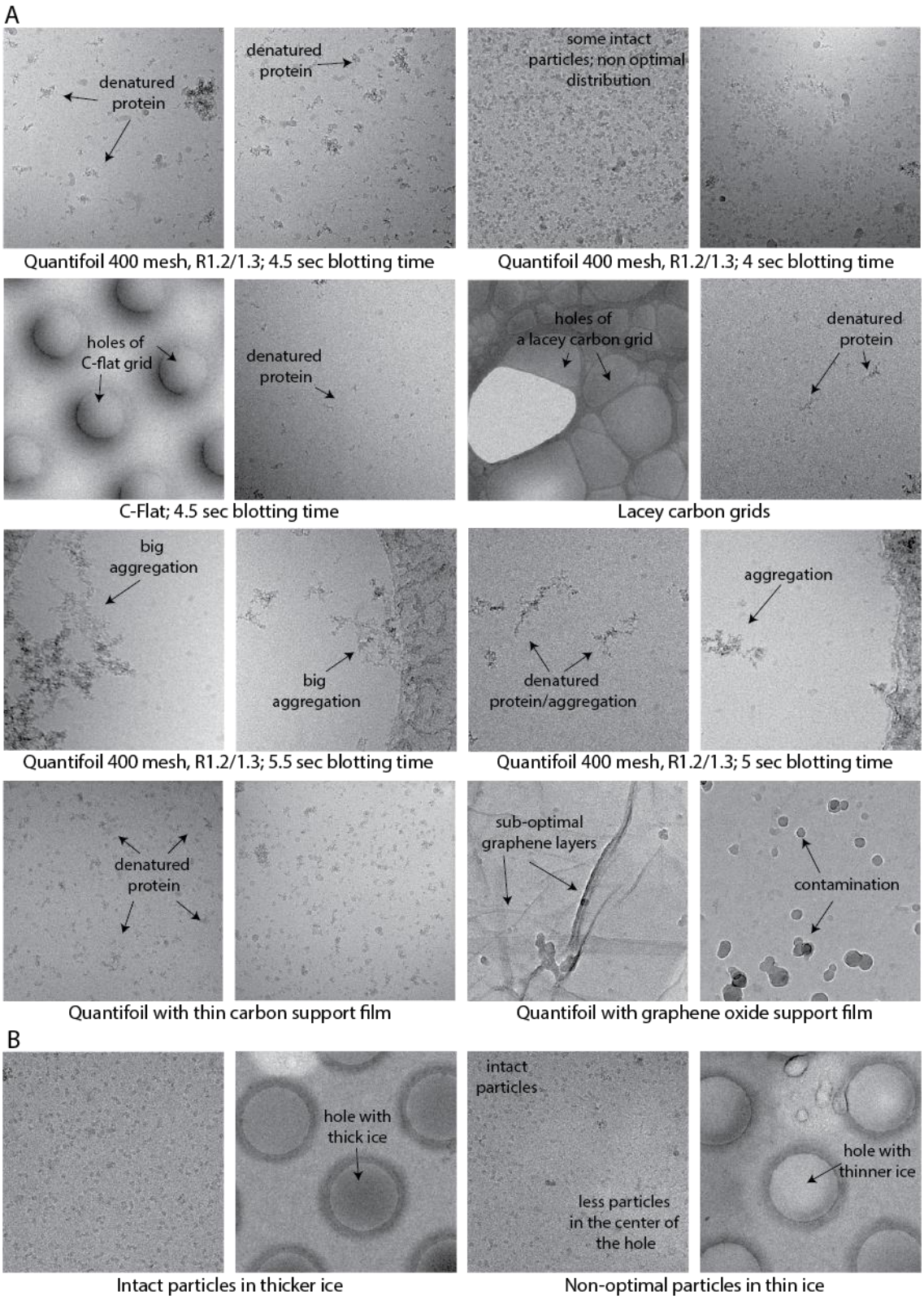

Figure 17: CryoEM analysis of CBF3 core complex

A Images of either empty holes or holes exhibiting denatured protein of grids prepared with a Vitrobot. **B** Successful freezing conditions of core complex using a manual plunge freezer and open-hole Quantifoil grids. The left panel shows nice particles in thick ice and an

overview of those holes in a lower magnification. The right panel shows micrographs from an area with thinner ice, in which the particles are sub-optimal, pushed to the side and degraded in the centre.

To avoid this problem, both ultra-thin carbon and graphene oxide support on holey carbon grids was tested. In both cases, however, no improvement was seen. Nevertheless, intact core complex could be detected on open-hole grids frozen with a manual plunge freezer (Figure 17B). This home-made apparatus produced overall much thicker ice than could be achieved with the Vitrobot, which was thought to make the difference in particle quality. The final freezing condition, therefore, was 0.1mg/ml protein solution straight after size exclusion chromatography frozen with a manual plunge freezer in liquid ethane, after 30sec incubation time on glow-discharged 1.2/1.3 400mesh or 2/2 200mesh Quantifoil copper grids.

4.4 Structural insights into CBF3 core

After successful screening of freezing conditions, data was collected for high resolution structure determination with cryoEM. In total two datasets, both on Titan Krios microscopes with Gatan K2 cameras, were collected: the first, at the Electron Bio-Imaging Center (eBIC) at Diamond Light Source, and the second in-house, after the establishment of the Francis Crick Institute's cryoEM facility (for details of data collections see Table 7). The first dataset, collected at eBIC, resulted in a 5.7Å cryoEM map, using a final particle number of 72,771. As reducing the particle number further resulted in a sharp drop in resolution, it was believed to be able to achieve higher resolution, if particle number would be increased. Therefore, the second dataset was collected, resulting in an initial set of 3 million particles. After multiple rounds of 2D classification, good class averages were obtained with clear secondary structure visible. 3D reconstruction, using the low pass filtered crystal structure of Cep3, showed some orientation bias, however, a final map of 3.6Å overall resolution could be obtained (Figure 18A-F). 3D classification didn't show any significant conformational heterogeneity and resolution anisotropy was assessed using 3DFSC (Tan et al., 2017; Figure 18E). The full image processing workflow is represented in Figure 19.

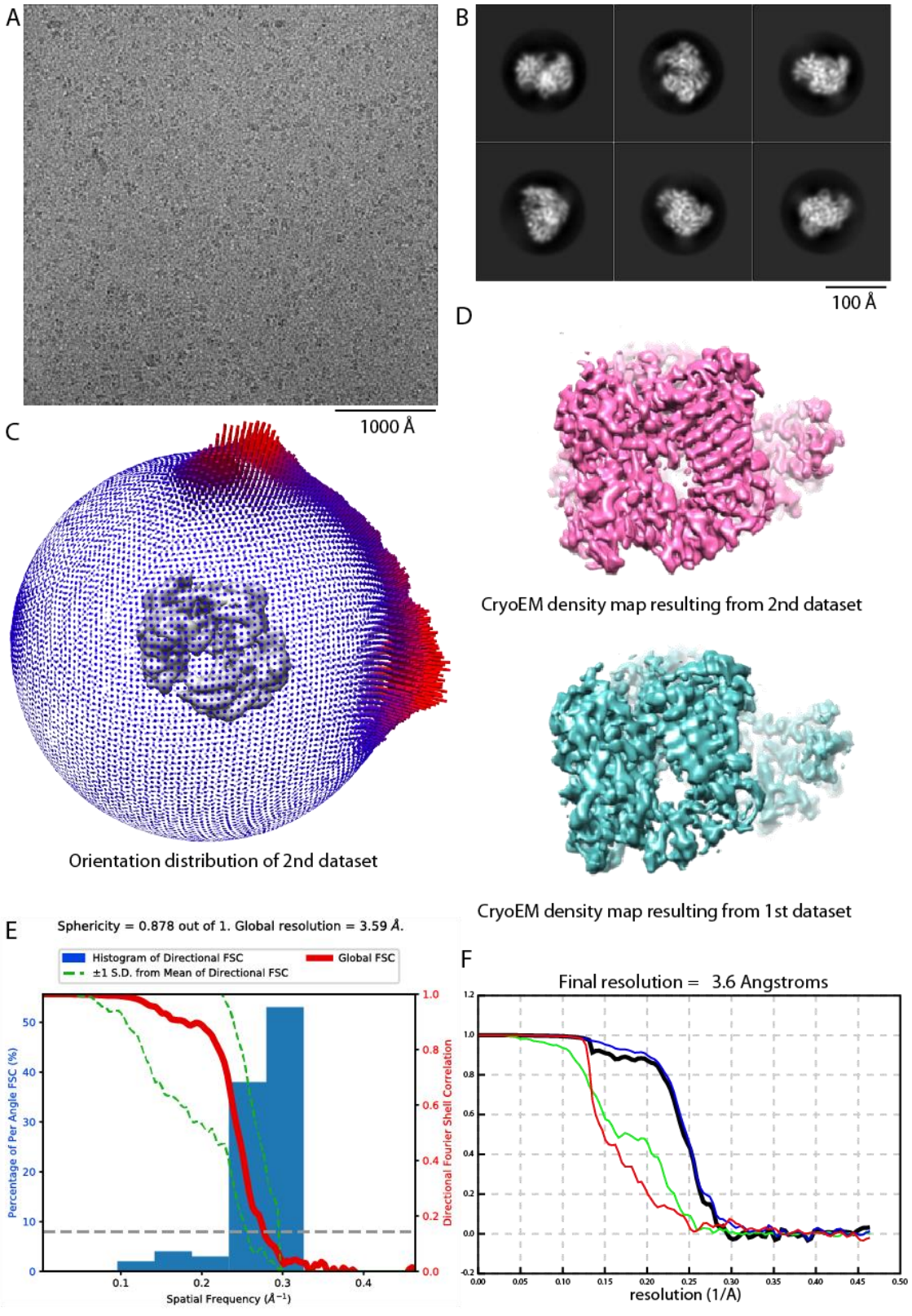

Figure 18: CryoEM data of CBF3 core complex

A An example micrograph recorded with a Titan Krios microscope and Gatan K2 camera. **B** Selected 2D classes of different orientations. **C** Plot of orientation distribution of particles

used for final reconstruction of the second dataset. **D** Comparison of the final cryoEM density map of the first and second dataset. **E** Assessment of resolution anisotropy in the final reconstruction of the second dataset calculated with 3DFSC (Tan et al., 2017). **F** Fourier shell correlation (FSC) curve, used to calculate the overall resolution limit of 3.6Å of the second dataset. Blue curve.. masked; green curve.. unmasked; black curve.. corrected and red curve.. phase randomised.

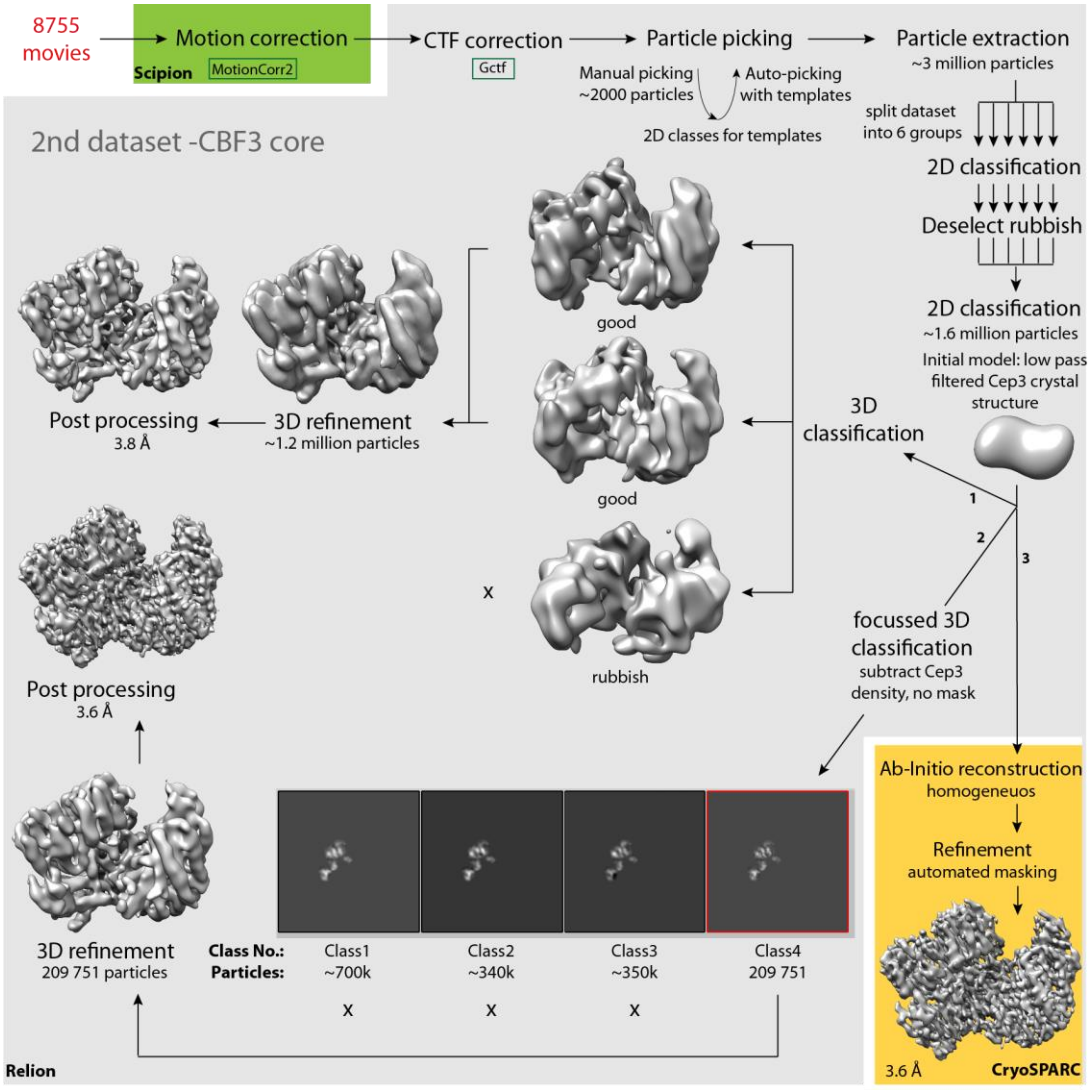

Figure 19: Processing workflow of the core wild type data

Overview of the image processing steps undertaken to achieve the 3.6Å map of the wild type core from the 2nd dataset collected in house.

4.4.1 Overall structure of CBF3 core

The final map allowed for easy docking of the two known crystal structures, Cep3Δ (2VEQ; N-terminal Gal4-domain deleted), and Skp1 (1NEX), leaving additional free density accounting for the unknown Ctf13 protein (Figure 20A, B and E). The latter forms a tightly associated dimer with Skp1, involving the Skp1 F-box binding C-terminus, as expected, and exhibits the typical arc-shape of a Leucine-rich-repeats (LRRs) protein (Figure 20C).

Together they bind on the outside lower edge of the Cep3 crescent, forming a central channel of about 30Å diameter in the overall shape (Figure 20B). The crystal structure of Skp1 fits directly into the EM density map, with two exceptions: a previously unresolved, extended loop (residues 105 to 112), which is in contact with Ctf13; and the C-terminal helices, which are involved in F-box interaction and exhibit a difference conformation in the cryoEM reconstruction. Interestingly, the binding of the Skp1/Ctf13 heterodimer stabilises the adjacent Gal4-domain of Cep3, whereas the second Gal4-domain is not resolved in the map, indicating that it remains flexible (Figure 20D). The rest of the Cep3 structure matches closely with the crystal structure, with another exception being a previously unresolved loop at the Ctf13 interaction surface. All these features are explained in more detail further below. Extra features of Cep3 and Skp1 proteins were built and sequence assigned and Cep3 and Skp1 models were refined using real-space refinement (Table 10). Deposition accession codes are EMD-4163 (EMDB database) and 6F07 (PDB).

Table 10: Refinement statistics for Cep3 and Skp1 structures

	Cep3	**Skp1**
CC_Mask	0.77	0.78
rmsd (bonds, Å)	0.01	0.01
rmsd (angles, °)	0.90	0.92
Rama favoured	94.3	91.9
Rama allowed	5.7	8.1
Rama outlier	0	0
Clash score	3.33	3.5
Molprobity	1.51	1.63

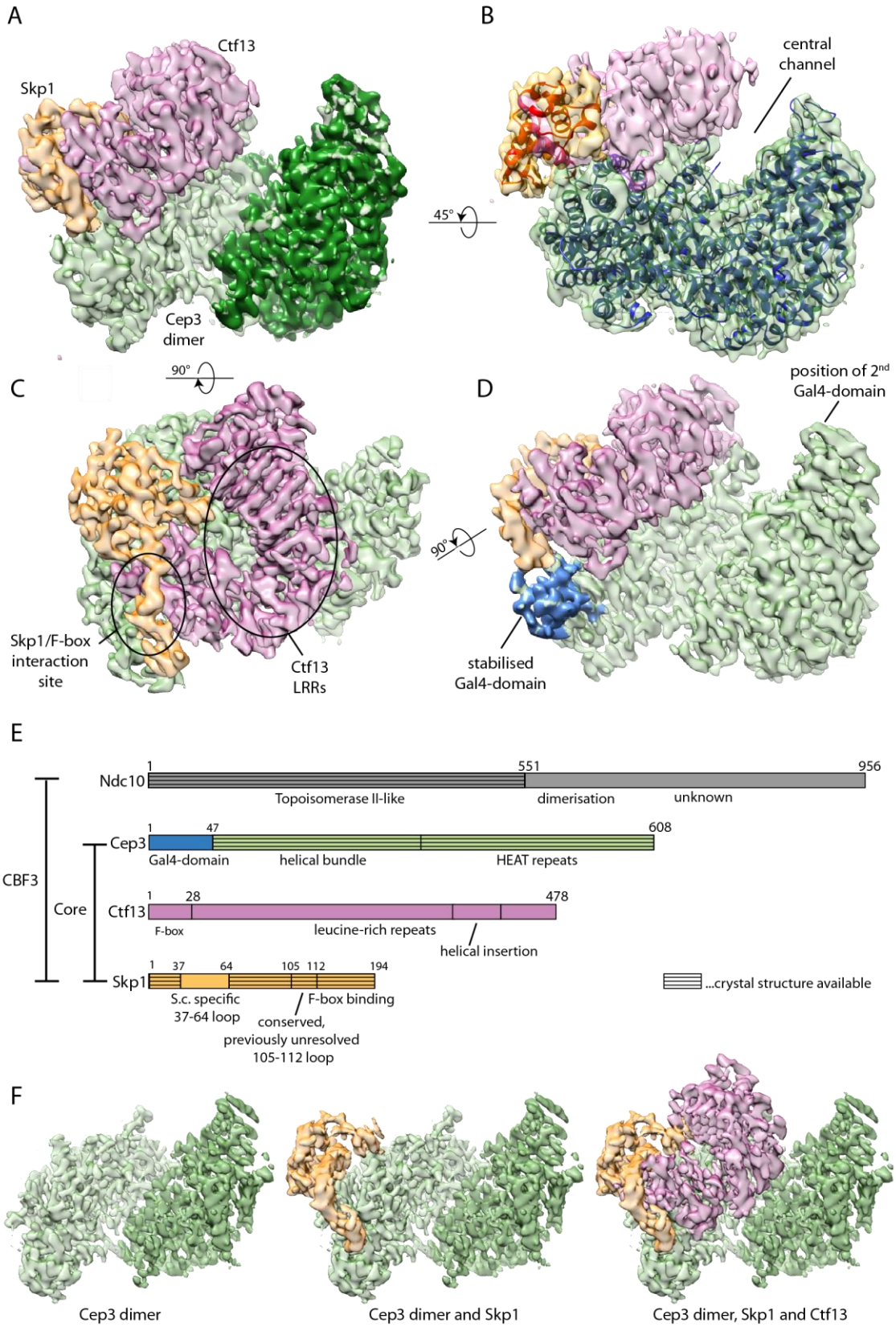

Figure 20: CBF3 core cryoEM map

A CryoEM density coloured according to subunits. **B** Coloured density with fitted crystal structures of Cep3∆ (blue; 2VEQ) and Skp1 (red; 1NEX). **C** Top view of density map, showing Skp1/Ctf13 heterodimer with overall features. **D** Side view showing the one stabilised Gal4-domain in blue and indicating the position of the second, still flexible and therefore unresolved Gal4-domain. **E** Schematic diagram of the four CBF3 subunits, colour coded as in the density maps with Ndc10 in grey. It is also indicated which parts of the subunits have been solved by crystallography before. **F** Dissected view of the core map to better illustrate how the individual subunits interact.

Note: Part of this figure was adjusted from Leber et al. 2018.

4.4.2 First structural insights into the subunit Ctf13

As described previously, Ctf13 on its own is a very unstable protein and eluded structural determination up to now. Our cryoEM structure, hence, provides the first insights into the shape of Ctf13: the N-terminal F-box forms a part of a small helical domain, which runs into seven LRRs, which are capped by a helical insertion distal to the F-box domain (Figure 21A and B). Ctf13 therefore, belongs to the family of F-box proteins containing LRRs (FBXLs), like Skp2.

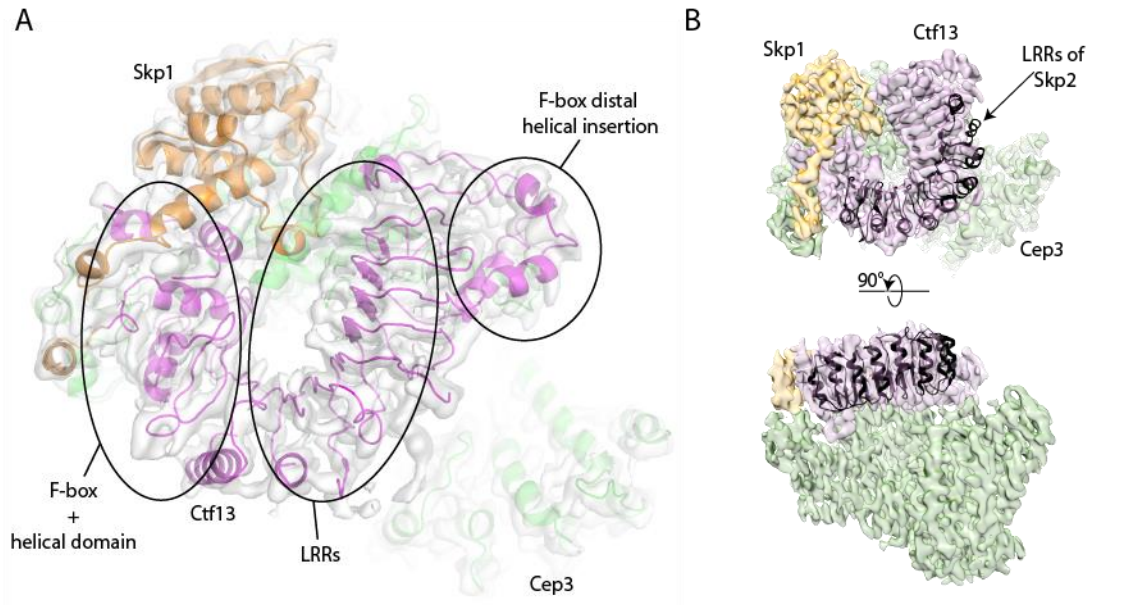

Figure 21: The structure of Ctf13

A Overview showing the three main elements of the Ctf13 structure: F-box domain, LRRs and helical insertion. **B** Colour coded cryoEM map in two views, overlaid with the LRRs from the Skp2 crystal structure (1FQV), to highlight the typical arc-shape of this domain.

Note: This figure was adjusted from Leber et al. 2018.

Despite clearly displaying this overall conformation, there is a high variability in the quality of the density map in this region. Only a few areas show clear side chain density, whereas the rest of the map does not. Even more critically, the connectivity between helices in the F-box domain, as well as in the external loops of the LRRs is unclear (Figure 22A-F), making reliable secondary structure and sequence assignment difficult, and a full atomic model could not be built. Instead a simple poly-alanine model was generated, with sequence only assigned to the N-terminal F-box, the helical insertion and the very C-terminus (Figure 23A and B).

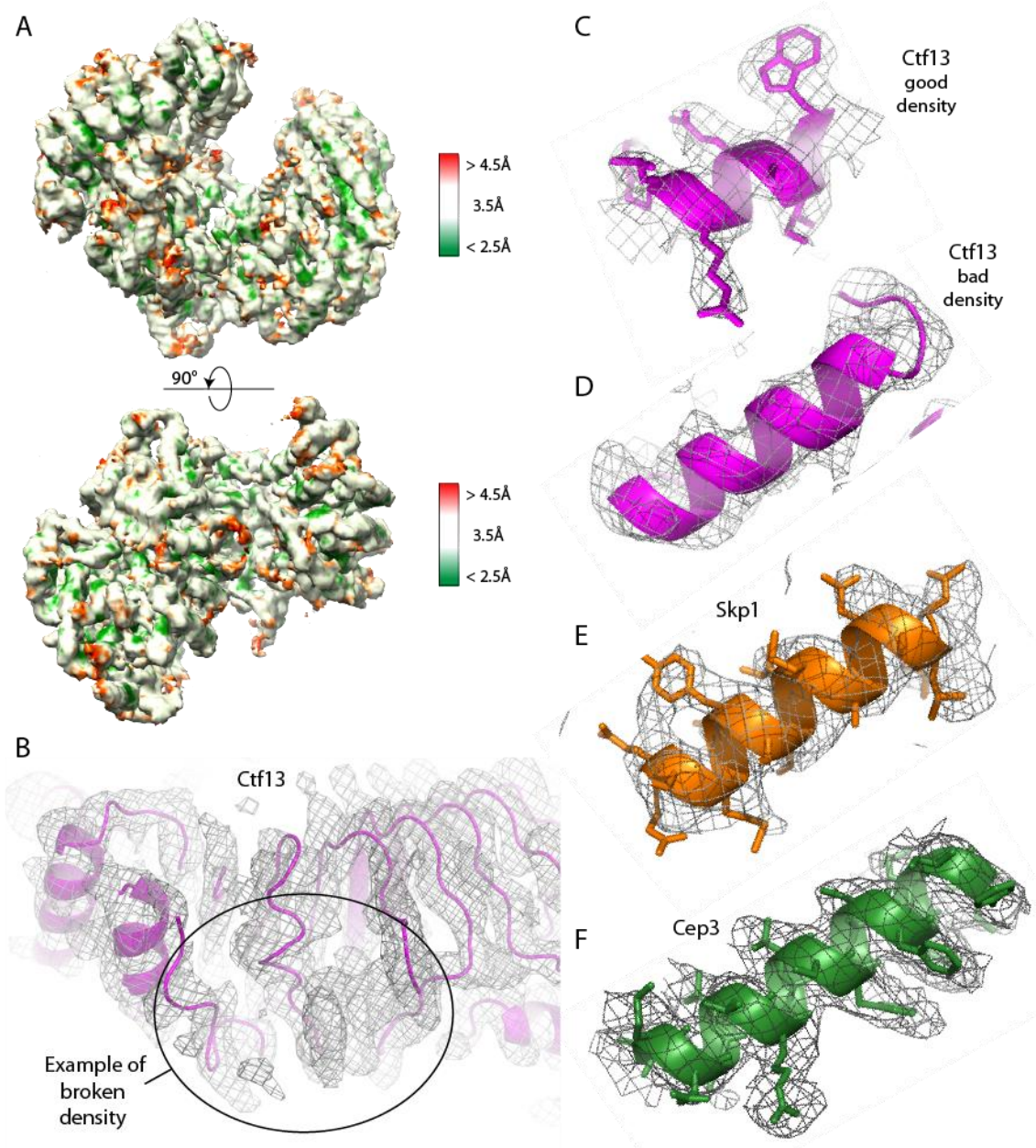

Figure 22: Local resolution and difference in quality of the cryoEM map

A EM density coloured after local resolution in two views. **B** Detailed view of broken density of Ctf13, making reliable tracing of the backbone impossible. **C** Example of a Ctf13 helix with clear side chain density. **D** In contrast, example of a Ctf13 helix with lower local resolution, where side chain density is lost completely. **E** Typical quality of density in the Skp1 subunit. **F** Typical density with visible side chains in the Cep3 heterodimer.
Note: This figure was adjusted from Leber et al. 2018.

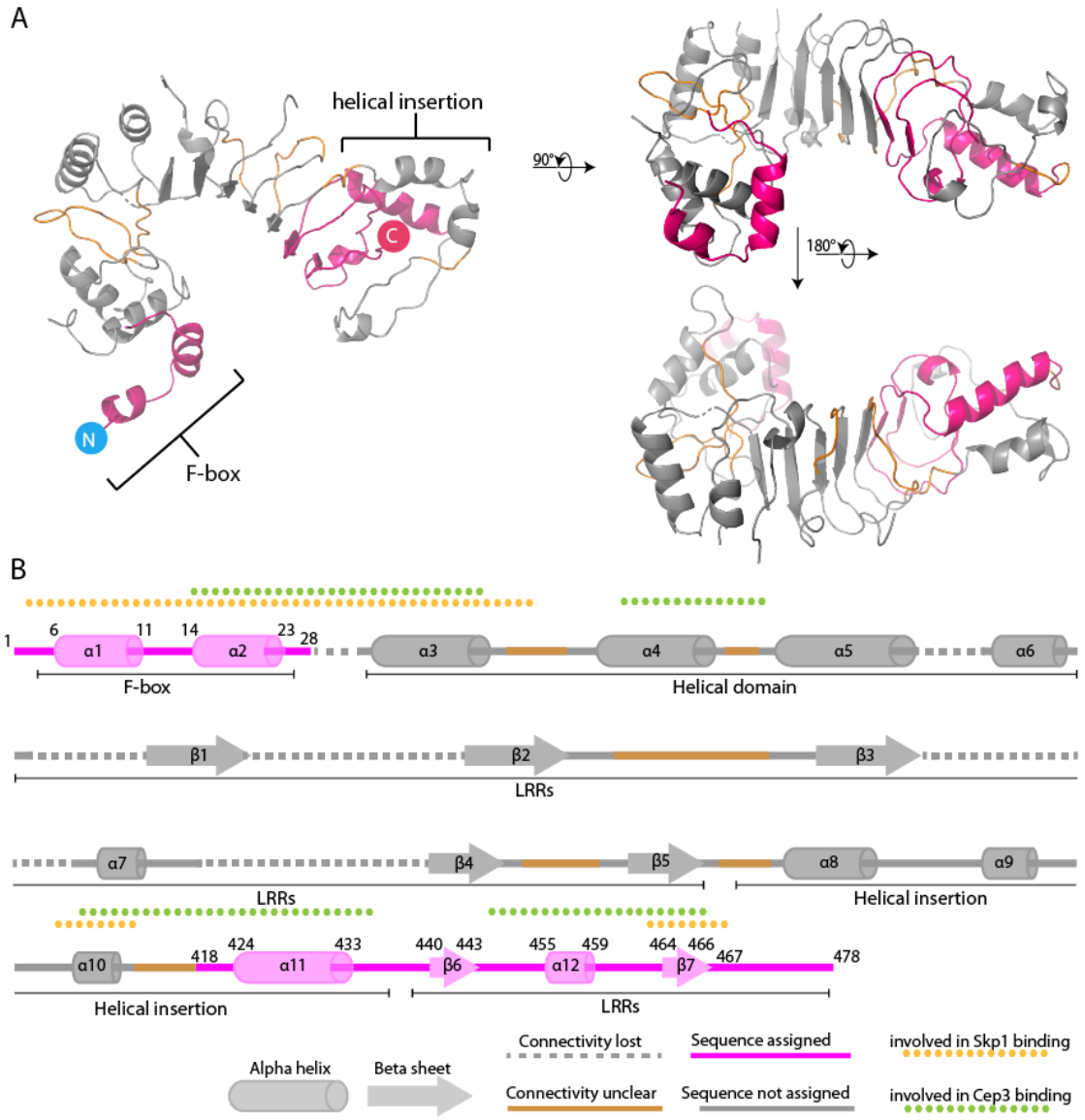

Figure 23: Incomplete model of Ctf13

A Ribbon diagram colour coded depending on sequence assignment (pink), poly-alanine chains (grey and brown) and areas where the connectivity is unclear (only brown). **B** Schematic representation of the Ctf13 model. *Note: This figure was adjusted from Leber et al. 2018.*

4.4.3 Skp1/F-box interaction

Skp1, as described in the introduction, is better known in the context of the SCF (Skp1/Cullin/F-box) ubiquitin ligases. There it acts as a linker between Cullin and a variety of substrate-binding F-box proteins. A number of crystal structures are available of Skp1 (both human and budding yeast) bound to different F-box proteins. While all of these structures exhibit the same Skp1/F-box binding mechanism, there is a clear difference in the core cryoEM map (Figure 24A-C). While the same three C-terminal helices of Skp1 (α6, α7 and α8) are involved, they show a more closed conformation, with the loop, connecting helices α7 and α8, shortened and folded. Equally to the crystal structures, the very N-terminus of the F-box of Ctf13 hooks underneath the Skp1 α6 helix. Helices α2 and α3, however, reveal a different orientation of the F-box. The distinction in this mode of binding might not be very surprising, taking in account the substantial sequence difference between the F-box of Ctf13 and the consensus F-box sequence (Figure 24D).

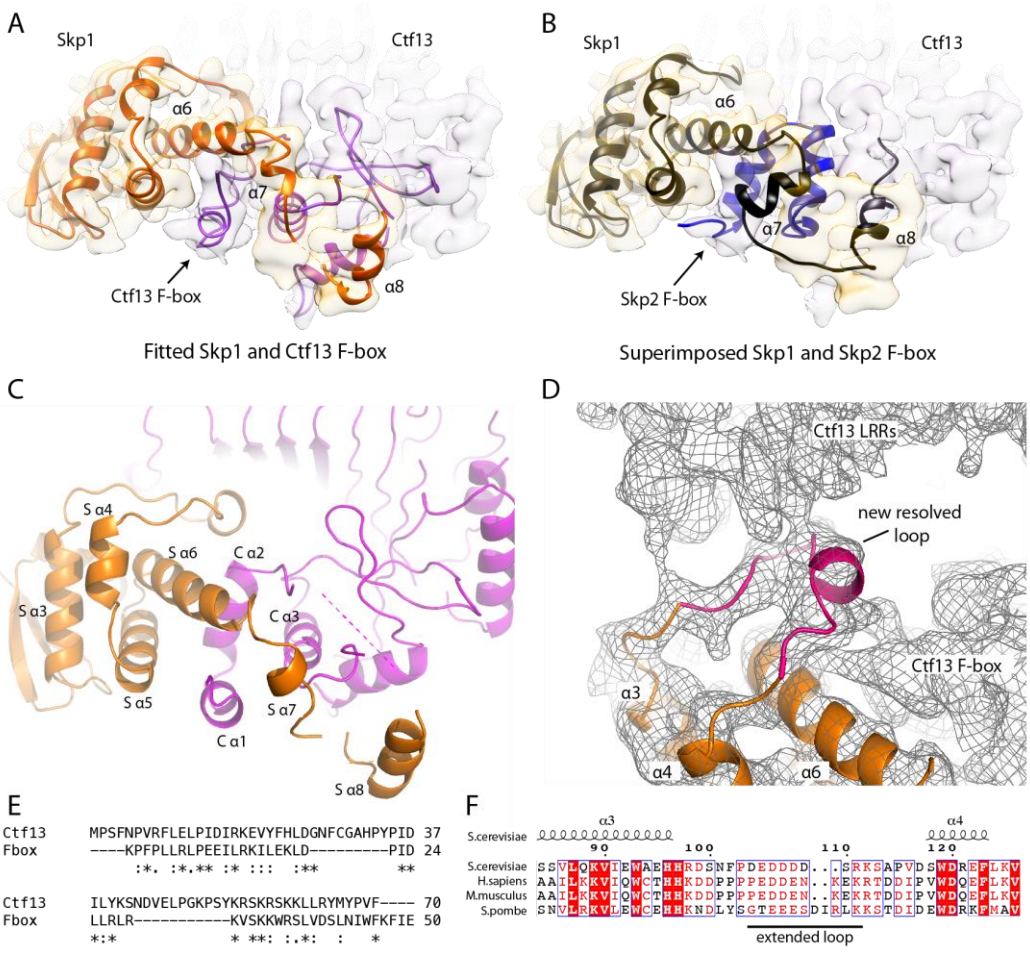

Figure 24: Skp1/F-box interaction

A-B Difference between the binding of Skp1 to the F-box of Ctf13 (**A**) and Skp2 (**B**). The cryoEM density (light yellow and pink) is overlaid with the models shown as ribbon diagrams. **C** Detailed view of Skp1 (orange) binding to the F-box of Ctf13 (pink). **D** Detailed view on the newly resolved 103 to 112 Skp1 loop contacting both the LRRs and the F-box of Ctf13. **E** Sequence alignment of Ctf13 with the F-box consensus sequence. **F** Sequence alignment of Skp1 from different species, highlighting the conserved 103-112 loop shown in **D**.
Note: Part of this figure was adjusted from Leber et al. 2018.

In addition to the differences in the F-box binding, a previously unresolved loop of Skp1 (residues 105-112) is structured in the cryoEM map. It contacts not only the LRRs of Ctf13 but surprisingly also the F-box. Although Ctf13 only exists in point centromere species, this loop is widely conserved (Figure 24E and F), suggesting there could be other F-box proteins bound in a similar fashion.

4.4.4 Interactions between Cep3, Ctf13 and Skp1

Besides giving insights into above discussed Ctf13 structure and new mode of F-box binding, the cryoEM map also provides first insights into how the three subunits interact. As already mentioned a heterodimer of Ctf13 and Skp1 binds to one of the two Cep3 monomers. Interestingly, the potential binding site of the other monomer remains free. The stoichiometry observed, therefore, agrees with the experimental data mentioned in 3.3. To investigate if a potential second Skp1/Ctf13 heterodimer would be able to bind this other site, the core density was superimposed with a mirror image of itself, aligning the two Cep3 densities. This showed a large steric clash between the two Ctf13 proteins, explaining the existing stoichiometry of 2x Cep3, 1x Skp1, 1x Ctf13 (Figure 25A and B).

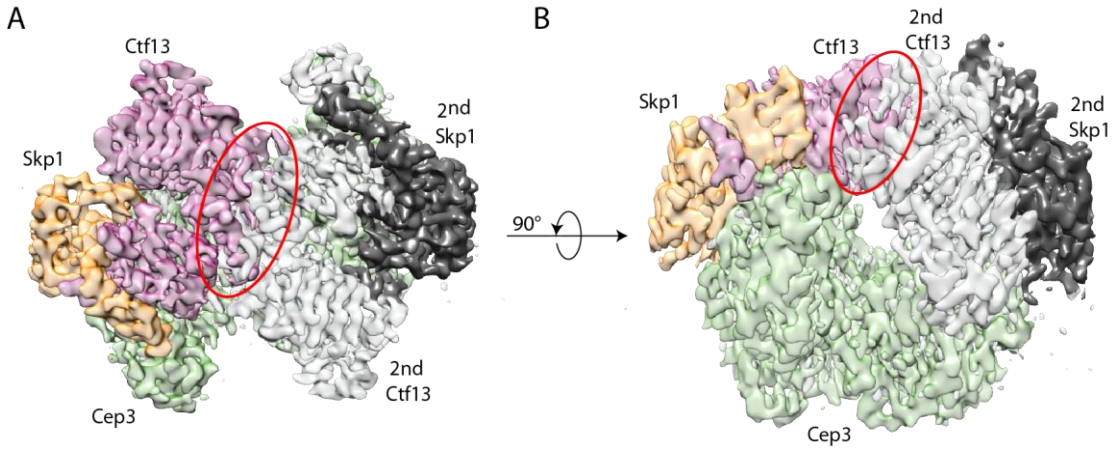

Figure 25: Structural rationalisation of stoichiometry

A-B Two views of the cryoEM density map with a hypothetical second Skp1/Ctf13 heterodimer (black/grey) superimposed on the 'unbound' Cep3 monomer, showing a steric clash (red ovals). *Note: This figure was adjusted from Leber et al. 2018.*

Interaction surfaces between Ctf13 and the other two subunits are vast and conform well with previous biochemical studies (Russell, 1999), which suggested that Ctf13 interacts with all other subunits and therefore forms the central part of CBF3. Whereas its interaction with Skp1 is through the N-terminal F-box and also via the C-terminal part of the LRRs, it binds to Cep3 via multiple surfaces spread over the entire length of the protein (Figure 26A and B). Consistently Russell, et al. 1999 showed that the two important regions of Ctf13 for the interaction with Skp1 are the N-terminal F-box and a C-terminal region, whereas most of the tested constructs showed some binding to Cep3, implying an extended binding interface.

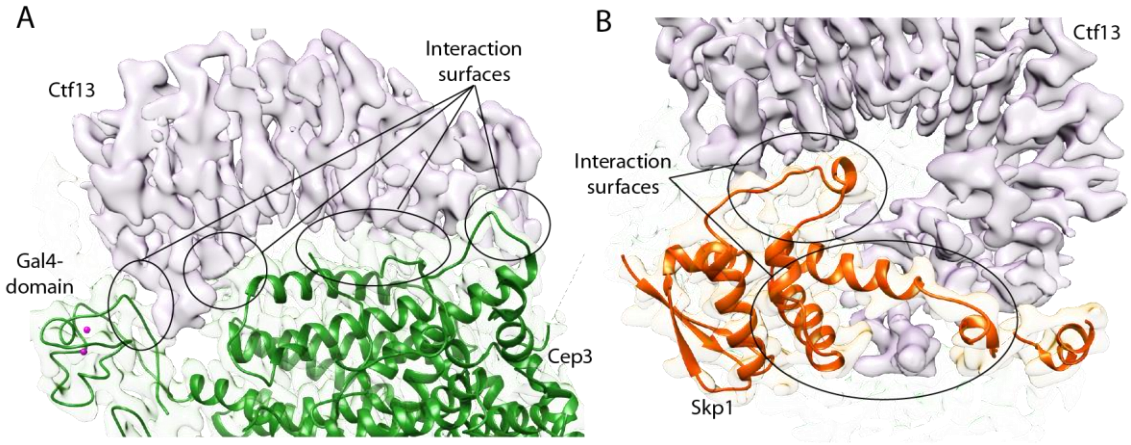

Figure 26: Interaction between the subunits

A Extensive interaction surfaces between Ctf13 and Cep3. **B** Two interaction sites between Ctf13 and Skp1. In both cases the EM density map is shown, colour-coded by subunit (Cep3.. light green, Skp1.. light yellow; Ctf13.. light pink), as well as the atomic models for Cep3 (green ribbon diagram) and Skp1 (red ribbon diagram). As no full atomic model could be built for Ctf13 only the EM density is shown. *Note: This figure was adjusted from Leber et al. 2018.*

Another notable trait of the core structure is that not only Ctf13 binds to both other subunits, but so does Skp1, placing it too into the centre of the CBF3 core structure. Interestingly, Skp1 binds to Cep3 through the same interface than it does to the SCF ligase subunit Cullin, highlighting additional similarities to this complex (Figure 27A-C; see 1.4.1). Indeed, mutant interaction studies could confirm involvement of the

same residues of Skp1 (N139 and Y140) shown to interrupt Skp1/Cullin interaction (Zheng et al., 2002). Co-expression and Strep pull-down from the lysate of Ndc10, Cep3, Ctf13 and a Skp1 construct with residues N139, Y140 mutated to two lysine (Skp1 NY_KK), showed that the mutated Skp1 cannot form a complex with the other subunits, and that it effectively gets replaced by endogenous Skp1. The wild-type HA-tagged Skp1, on the other hand, interacts with the others (Figure 27C and D).

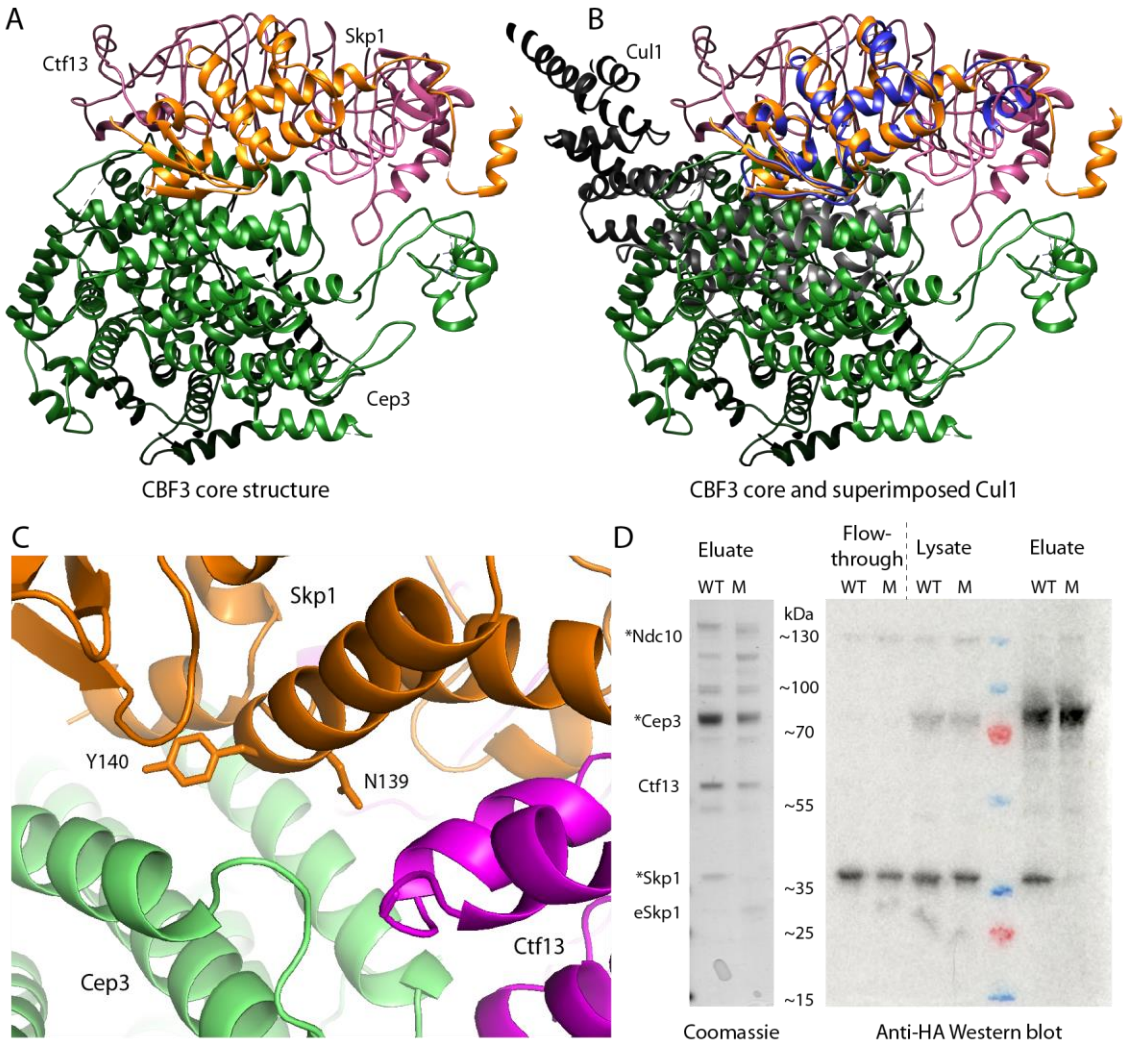

Figure 27: Similarity with interaction site of Skp1 to Cullin and Cep3

A Back-view of core structure with the Cep3 dimer in green, Skp1 in orange and Ctf13 in pink. **B** Same view of the core structure with Cul1 (dark grey) and Skp1 (blue) from the crystal structure of the SCF ligase (PDB: 1LDK) showing a vast sterical clash between Cep3 and Cul1, emphasising that both Cep3 and Cul1 bind at the same interface to Skp1. **C** Detailed view of the two Skp1 residues (N139, Y140) which are important in the interaction between Skp1 and Cullin (Zheng et al., 2002) as well as Cep3. **D** Coomassie stain (left panel) of

eluates and western blot (right panel) of flow-through, lysate and eluates showing that only HA-tagged wild-type (WT) Skp1, but not the N139, Y140 mutant (M) can form a complex. Mutant Skp1 gets replaced by endogenous Skp1 (eSkp1), which is migrating faster due to the lack of HA-tag. The asterisks indicate HA-tagged proteins. *Note: Part of this figure was adjusted from Leber et al. 2018.*

As already mentioned before, the Gal4-domain adjacent to the Skp1/Ctf13 heterodimer is stabilised and therefore resolved in the cryoEM map. The density was of sufficient quality to visualise the Zn_2Cys_6 cluster, trace the connecting loop and assign the correct amino acid sequence (Figure 28A and B).

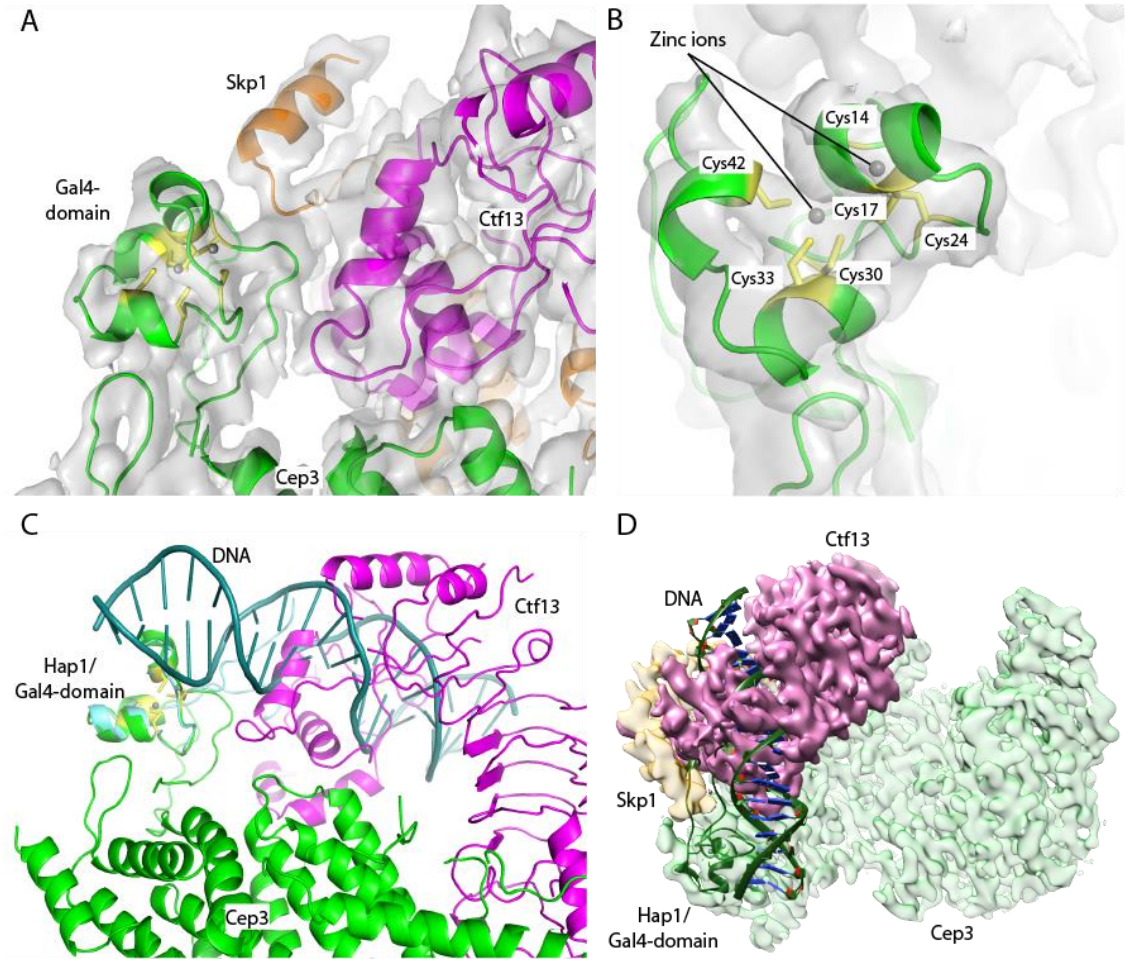

Figure 28: The stabilised Gal4-domain and its influence on DNA binding

A Overview of the stabilised Gal4-domain, showing the cryoEM density map in grey, and the ribbon diagrams of Ctf13 (pink), Skp1(orange) and Cep3 (green). The 6 cysteines are highlighted in yellow. **B** Detailed top view, highlighting the positions of the zinc ions. **C** Detailed view of superposition of Hap1 (blue) bound to DNA on the stabilised Gal4-domain,

showing a steric clash between DNA and Ctf13. **D** Overview of the previous showing the whole core cryoEM map. *Note: This figure was adjusted from Leber et al. 2018.*

Both Skp1 and Ctf13 contact the Gal4-domain, which must account for this stabilising effect. The second Gal4-domain, on the other hand, is not resolved in the map, which is not surprising given it is connected to the rest of the protein through a rather long (14 residues), flexible linker. To examine how this interesting, novel feature might influence the DNA binding ability of CBF3 core, the crystal structure of the homologous Hap1 domain bound to DNA (King et al., 1999) was superimposed on the stabilised Gal4-domain. This allows to trace the likely path of DNA and shows quite clearly a severe clash between the DNA and the LRRs of Ctf13 (Figure 28C and D). Skp1/Ctf13 therefore seem to lock this Gal4-domain in an inactive position, leaving the second still flexible Gal4-domain to bind to the CCG triplet in the centromere.

Finally, a conserved loop of Cep3 (residues 303-337) is stabilised in the cryoEM map, whereas it is unresolved in the crystal structure of Cep3Δ. Moreover, it is only structured on the monomer bound by Skp1/Ctf13 and not the other monomer. Both these findings suggest that this loop might be involved in the interaction between Cep3 and Ctf13 and therefore is getting stabilised upon binding (Figure 29A). Mutation studies of conserved residues of this loop (residues L319, L320 and Y325, Table 8), however, showed no influence on complex formation (Figure 29B and C) as determined with recombinant expression of the mutants and pull-down studies from the lysate. Given the vast interaction surface between Cep3 and Ctf13, shown in Figure 26A, this might not be surprising. It might not even exclude that these residues are important for the interaction between Cep3 and the Skp1/CTf13 heterodimer.

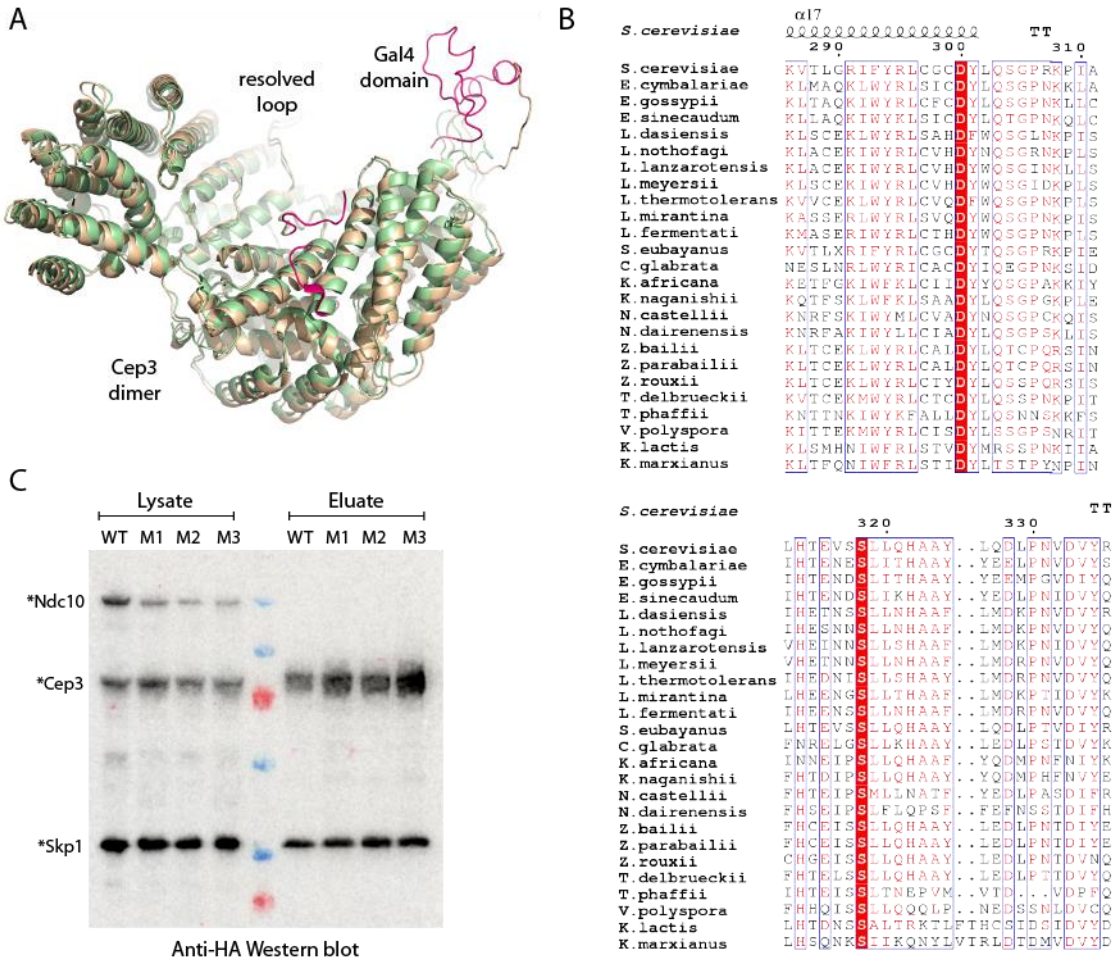

Figure 29: Mutational studies of the Cep3 303-337 loop

A Comparison of the crystal (green) and cryoEM (brown) structure of Cep3. The newly resolved 303-337 loop and Gal4-domain of the cryoEM structure are coloured in magenta. **B** Sequence alignment of Cep3 homologues. Red box: fully conserved residues; red font: conserved residues. **C** Western blot with Ha-tag antibody of lysate and streptavidin pull down eluate from cells co-expressing all four CBF3 subunits, with either wild-type or mutant Cep3. Mutations in Cep3 were L319 and L320 to serine and alanine, respectively (M1), Y325 to serine (M2) or a combination of both (M3). Mutations did not interrupt complex formation.

Note: Part of this figure was adjusted from Leber et al. 2018.

Chapter 5. Results 3: Functional studies of CBF3

The cryoEM structure discussed in the previous chapter gave rise to many questions about functionality of the core as well as the full complex. Biochemical studies were conducted to better understand how CBF3 binds to the centromere and how it might get assembled. It was also attempted to find conditions in which structural studies of the full or core complex bound to the centromere would be feasible.

5.1 DNA-binding of CBF3

CBF3 binds sequence-specifically to the CDEIII of the point centromere. Both Ndc10 and Cep3 have been shown to exhibit DNA binding. The former, however, binds sequence non-specifically and it is the Gal4-domains of the Cep3 homodimer, which account for the specificity (see 1.4.1). The canonical recognition sequence for these motifs is a CCG/GGC triplet, and as all known Gal4-domain containing transcription factors are dimers (see 1.4.2), they bind to a site containing either a direct or inverted repeat of this sequence. However, there is only one conserved CCG/GGC triplet in the point centromere. A recurrent question in the field therefore is, how CBF3, having two copies of the Gal4-domain, is binding to a single, asymmetric site on the centromere. Also, it has been suggested that to be able to bind to the centromere all four subunits have to be present (Sorger et al., 1995, Stemmann and Lechner, 1996, Kaplan, 1997, Espelin, 1997, Pietrasanta, 1999), which given the fact that both Ndc10 and Cep3 can bind DNA on their own is surprising (Bellizzi, 2007, Purvis and Singleton, 2008, Cho and Harrison, 2012, Perriches and Singleton, 2012).

5.1.1 Dephosphorylated core can bind the centromere on its own

To test if CBF3 core complex can bind to centromeric DNA without Ndc10 present, EMSA studies were conducted. At first, in agreement with above mentioned studies no binding activity of the core complex could be detected. However, in an attempt to test different hypothesis about complex formation and activity, it was found that core complex can bind centromeric DNA, but only if pre-treated with Lambda Protein Phosphatase (Figure 30A). This surprising and new finding suggests a possible regulation of DNA-binding activity through phosphorylation. Previous studies of CBF3, however, have only pinpointed phosphorylation important for complex

assembly. One should note that the core cryoEM structure presented in this thesis represents the DNA-binding inactive state of the core complex.

In agreement with previous results, DNA binding is facilitated by the Gal4-domains, as a construct lacking this domain shows no DNA binding activity with or without Lambda phosphatase treatment (Figure 30A). It was also confirmed that the binding is sequence-specific, as shown with a DNA-competition experiment. DNA with the CCG/GGC triplet mutated, failed to compete out a labelled wild-type CEN3 probe, underlining its importance for DNA-binding. Wild-type or DNA with the potential second TGT triplet mutated, on the other hand, was able to compete out the probe (Figure 30B).

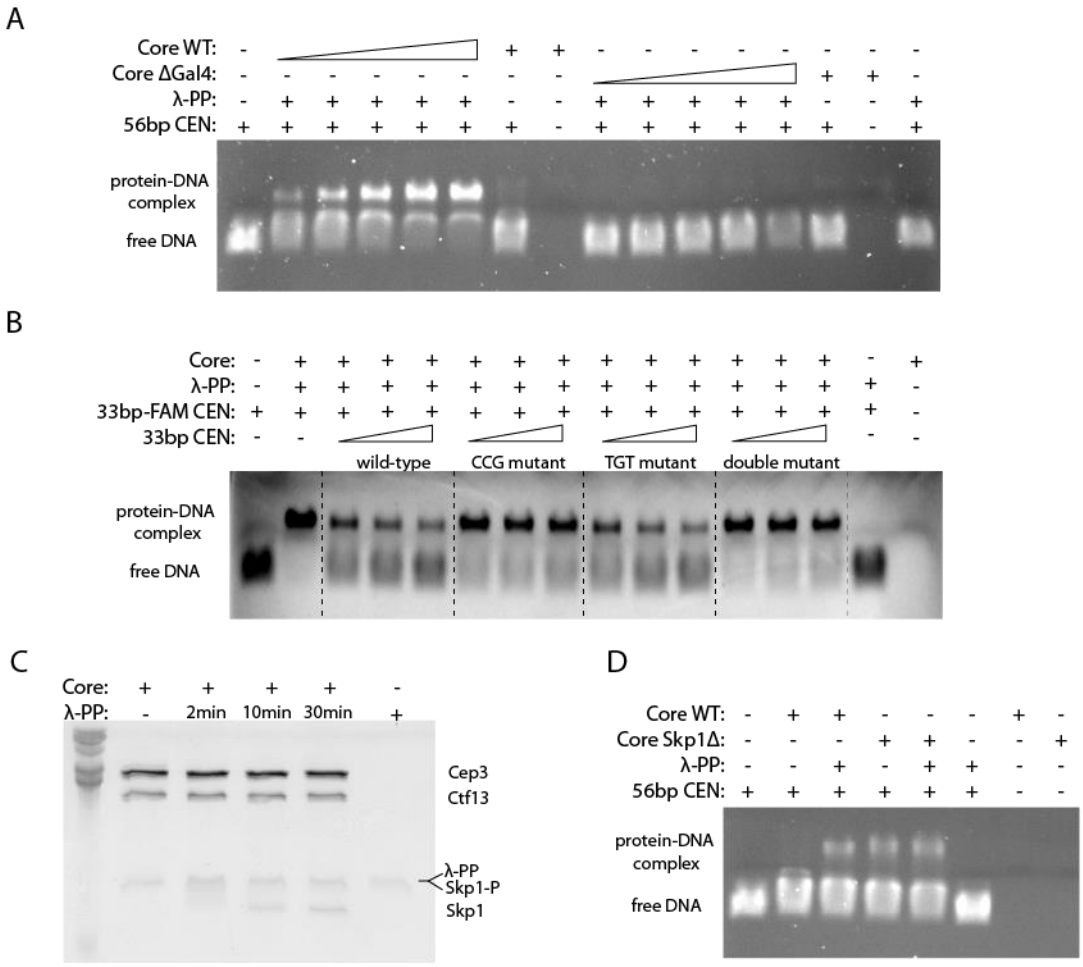

Figure 30: DNA-binding studies

A EMSA comparing wild-type core (CoreWT) with core lacking the N-terminal Gal4 domain (CoreΔGal4). Controls include free DNA (lane), reactions without λ-phosphatase (λ-PP) treatment (lane7 and 14), protein alone (lane 8 and 15) and DNA alone with λ-PP (lane 16).

B EMSA to test for sequence specificity. Labelled centromeric DNA was competed with wild-type, a CCG mutant, a TGT mutant and a double mutant. See table XX for sequences of DNA used. C Phos-tag SDS-PAGE/Coomassie stain of core complex without and with increasing amounts of time of λ-PP treatment. D EMSA demonstrating the λ-PP treatment dependency of core complex for DNA binding, and the lack of the same for a construct lacking an N-terminal Skp1 loop (residues 37-64). *Note: This figure was adjusted from Leber et al. 2018.*

Mass spectrometry analysis, carried out by the Francis Crick's Proteomics STP, found some evidence of phosphorylation on all four subunits. To identify which subunit is dephosphorylated by Lambda phosphatase, Phos-tag™ SDS-PAGE was used. The Phos-tag™ reagent produces a shift of phosphorylated proteins in the gel. Untreated sample and Lambda phosphatase treated sample was run on the gel, and a clear shift was observed for Skp1, whilst the Cep3 and Ctf13 bands were unchanged (Figure 30C). It should be noted, that no shift of a band does not necessary mean that this protein is not phosphorylated at all, but only that no dephosphorylation takes place. Treatment with calf intestine phosphatase (CIP) for example did not lead to any shifts. Previous studies identified a number of phosphorylation sites on Skp1 (see 1.4.1), including four phosphorylations on a budding yeast-specific loop between residues 37 to 64. This loop is not resolved in the cryoEM map, suggesting it is extended and flexible and indeed one can express a construct lacking this loop, as has been done for crystallisation of Skp1 (Orlicky et al., 2003, Tang et al., 2012, Willhoft et al., 2017). To test if the phosphorylation status of this loop is indeed responsible for the difference in DNA binding activity, core complex with Skp1 lacking the loop (coreSΔ) was expressed, purified and tested with EMSA. Whilst wild-type core showed a clear dependency on Lambda phosphatase treatment, coreSΔ was able to bind to DNA with and without being dephosphorylated (Figure 30D), confirming that the DNA binding activity of the core complex *in vitro* is regulated by the phosphorylation status of this budding yeast-specific loop. In-house intact mass analysis, carried out by the Proteomics STP, could correspondingly confirm four phosphorylations on Skp1 (Figure 31).

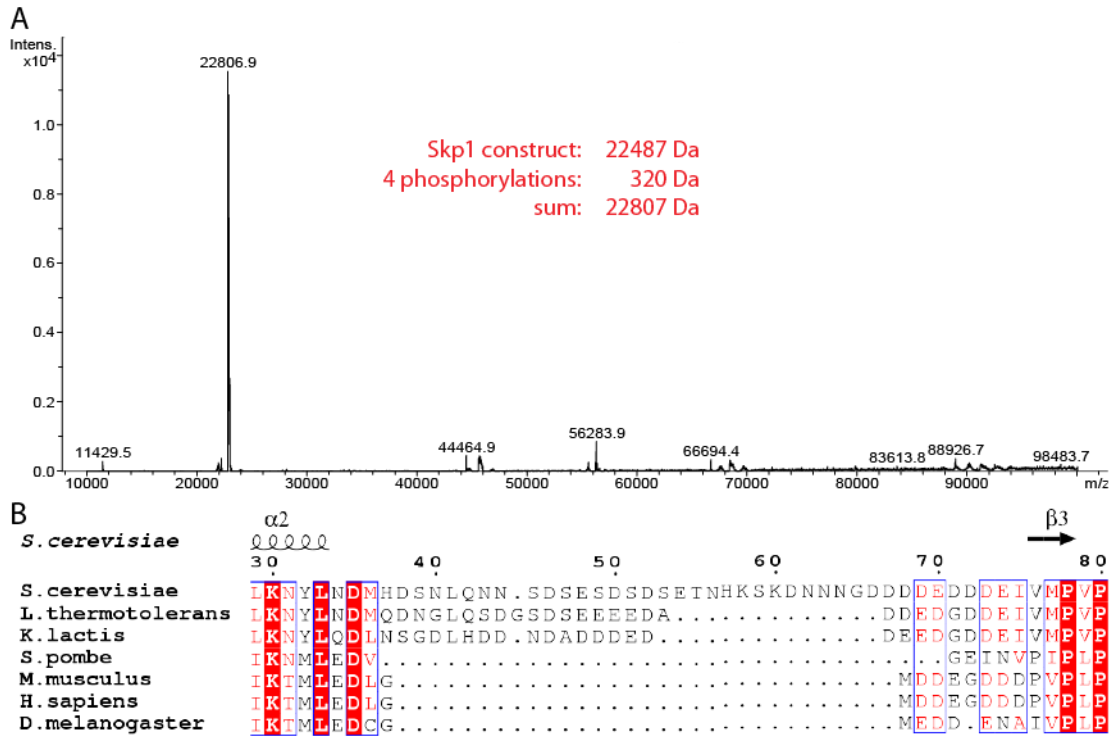

Figure 31: Mass spectroscopy analysis of Skp1

A Intact mass analysis of CBF3 core, confirming the presence of four phosphorylations. **B** Sequence alignment of Skp1 37-64 loop.

Note: Intact mass analysis was solely done by the Francis Crick Institute's Proteomics STP

5.1.2 DNA binding of full CBF3 still influenced by phosphorylation status

Next it was tested if the full complex exhibits the same behaviour as the core in EMSA studies. As Ndc10 alone also binds to DNA, it was expected to see some binding even without Lambda phosphatase treatment. The binding after treatment should be stronger, however, if there is the same regulatory mechanism. As expected, EMSA analysis showed that the full complex can bind both with and without phosphatase treatment, the dephosphorylated CBF3, however, bound more strongly (Figure 32).

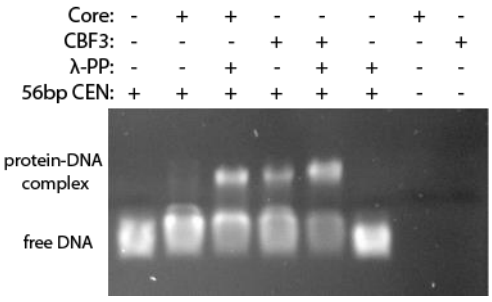

Figure 32: DNA-binding of the full CBF3 complex

EMSA analysis of CBF3 full (CBF3) and core complex (core), both with and without λ-protein phosphatase (λ-PP) treatment. DNA alone (lane 1), DNA and phosphatase (lane 6) and both protein complexes alone (lane 7 and 8) were run as controls.

5.1.3 Cell cycle-dependency of Skp1 phosphorylation

To test if the there is a cell-cycle dependent difference in the phosphorylation status of the 37-64 loop of Skp1, strains with tagged endogenous genes of Skp1 and Cep3 were constructed. Cep3 was C-terminally tagged to allow for CoIP to separate Skp1 bound within the CBF3 complex from free Skp1, as well as Skp1 included in SCF ligases and the RAVE complex. Skp1 was HA-tagged to allow for western blot detection. A strain with the Skp1 37-64 loop deleted (Skp1Δ) was prepared as control to ensure any phosphorylation seen is indeed positioned on this loop. Both strains progressed through the cell cycle identically, as shown with FACS analysis (Figure 33A). No change in phosphorylation or Skp1 protein amount could be detected in the Skp1 wild-type strain (Figure 33B).

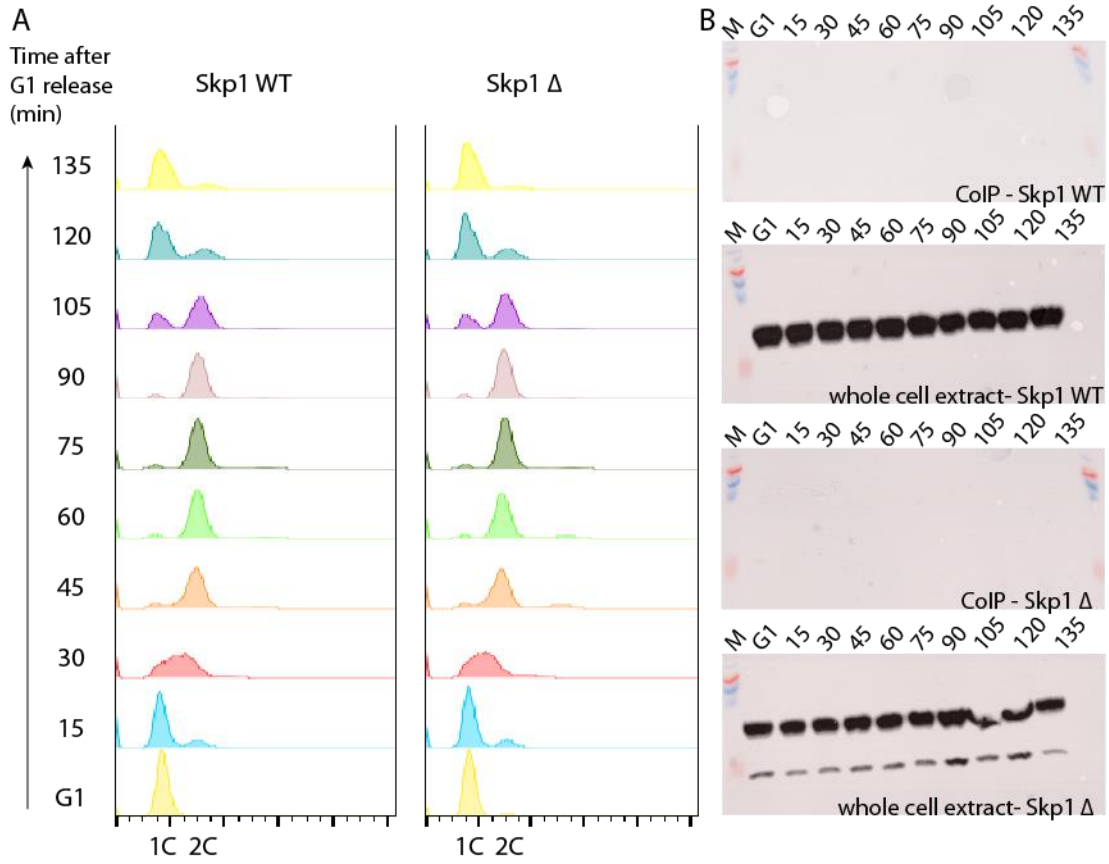

Figure 33: Cell-cycle dependency of Skp1 phosphorylation

G1 arrested cells of HA-tagged Skp1 wild-type (WT) or 37-64 loop deleted (Δ) were released into a synchronous cell cycle progression, which was monitored using FACS analysis of DNA content (**A**). Phosphorylation status of either Skp1 WT or Δ was monitored using Phos-tag SDS-PAGE and western blotting against HA tag (**B**). Both the whole cell extracts as well as

the bound fractions of the CoIP were tested. For the CoIP a PK6-tag was fused on the C-terminus of Cep3, to separate Skp1 bound within the CBF3 complex, from other Skp1 in the cell. Surprisingly two bands were seen in the whole cell extract of the Skp1Δ strain, the upper corresponding to wild-type Skp1 and the lower possibly to Skp1Δ. Replacing endogenous Skp1 with tagged Skp1Δ therefore seemed to have failed, although genotyping showed a positive result. Reconstruction of this strain could not be repeated due to time constraints. However, no difference in total wild-type Skp1 phosphorylation could be detected. Furthermore, the CoIP of both wild-type and Skp1Δ did not yield any detectable Skp1, possibly due to the cell lysis method used (see 7.3).

5.1.4 *In vitro* reconstitution of a CBF3/DNA complex

Given the successful demonstration that both the core and full complex can bind to the centromeric DNA, it was attempted to reconstitute a protein/DNA complex for structural studies by cryoEM. This was done both with wild-type CBF3/core and CBF3/core containing Skp1Δ, the DNA binding 'active' state.

First, in an attempt to purify excess DNA and unbound protein from protein/DNA complexes, analytical size exclusion was tested. Protein was mixed with an excess of centromeric DNA and injected into a Superose 6 column. All tested CBF3 complexes, however, eluted separately from the DNA, independent from buffer conditions, or DNA length (Figure 34A and B).

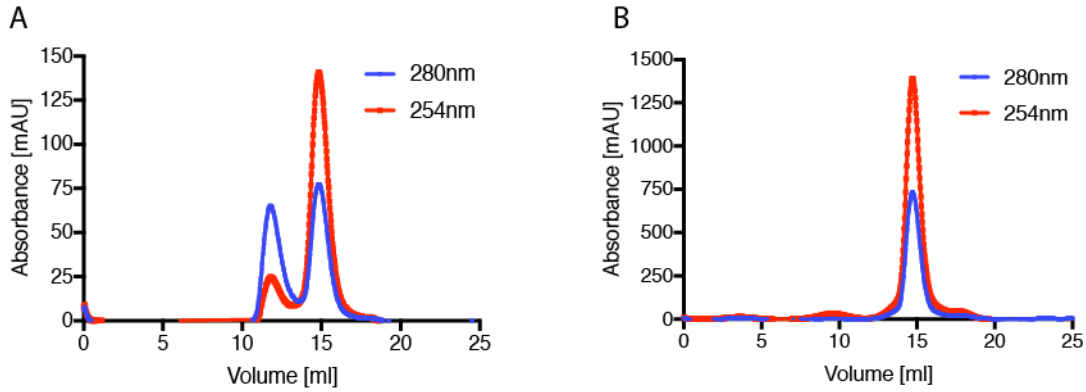

Figure 34: Size exclusion chromatography to test protein/DNA formation

A Profile of CBF3 SΔ mixed with a 1:2 excess of 56bp CEN3 dsDNA. The first peak corresponds to CBF3 without bound DNA and the second peak to DNA alone. **B** Profile of 56bp CEN3 dsDNA run on the same column and buffer alone. For both chromatographies a

Superose 6 10/300 column in 10mM HEPES, 300mM NaCl and 1mM DTT was used. Chromatographies with other constructs/DNA fragments behaved identical (data not shown).

As size exclusion seemed to be too harsh to obtain a protein/DNA complex, it was attempted to simply add an equimolar amount of centromeric DNA to the protein before making negative-stain grids. However, no difference in particles could be seen (Figure 35A-E). It should be noted that, although DNA is not easily visible in negative stain EM, one would expect to see a difference at least with the full CBF3 complex. That is, as unbound CBF3 seems to exhibit unstructured Ndc10 forming indistinct density and core particles, and DNA is thought to stabilise the full complex and lead to a more structured Ndc10. Furthermore, grids with longer pieces of DNA, which can be easily seen in the micrographs, also provided no difference in protein particles (Figure 35C). Although DNA-binding EMSAs showed that the core can bind DNA at a salt concentration of 300mM, which is needed to prevent protein from aggregating, it was tested if lowering the salt will help formation of protein/DNA complexes. This was especially important, as salt concentration will rise significantly during the 1min incubation time of the sample on the grid due to evaporation. This increased salt concentration could inhibit DNA binding. As CBF3 core aggregates if kept in lower salt buffer, it was added to the grid in normal conditions, incubated to allow protein particles to adhere to the grid and subsequently it was washed twice with a low salt buffer (10mM HEPES, pH 7.5, 150mM NaCl, 1mM DTT) containing 3ug/ml 56bp CEN3 dsDNA. That way aggregation was hindered, as particles were already stuck to the carbon. Indeed, screening of grids showed that particles did not aggregate. They, however, still had no DNA bound and looked identical to unbound particles (Figure 35F).

As already mentioned, negative stain is not optimal to visualise DNA, particularly if it is short. Therefore, cryoEM grids of coreSΔ (DNA-binding active) with a slight excess of DNA were prepared and a dataset was collected to test if DNA density can be visualised in the cryoEM reconstruction. Besides this dataset was also recorded to see if this core structure, with the Skp1 37-64 loop deleted, exhibits a different conformation compared to the wild-type core structure described in 4.4. Such a difference could give insights into why this construct is DNA binding active, even if no DNA density can be seen. The results are described in the following section.

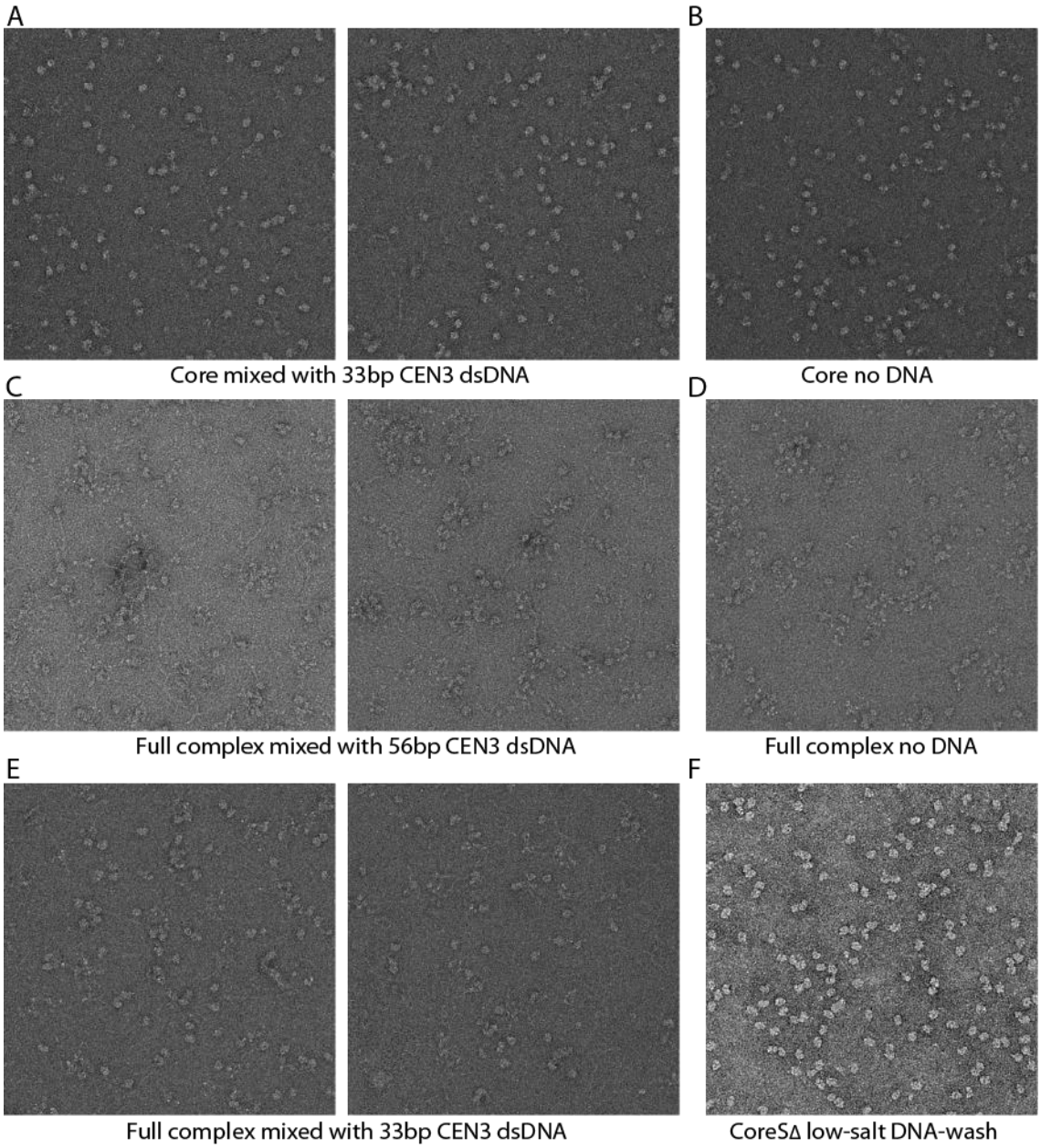

Figure 35: Negative stain analysis of CBF3 and core mixed with DNA

Protein was prepared in 300mM salt buffer and mixed with DNA in a molar ratio of 1:1.2. Core was prepared with 33bp DNA (A) and without (B) and full complex was tested with 56bp DNA (C), without any DNA (D) or with 33bp DNA (E). F CoreSΔ was added to the grids and subsequently washed with a 150mM salt buffer containing 56bp CEN3 dsDNA to test if lowering the salt will help DNA binding.

5.1.5 Structure of coreSΔ

A dataset was collected on the in-house Titan Krios and K2 camera of grids containing coreSΔ with added DNA and processed with Relion (see Table 7 for data collection parameters and Figure 36 for processing workflow).

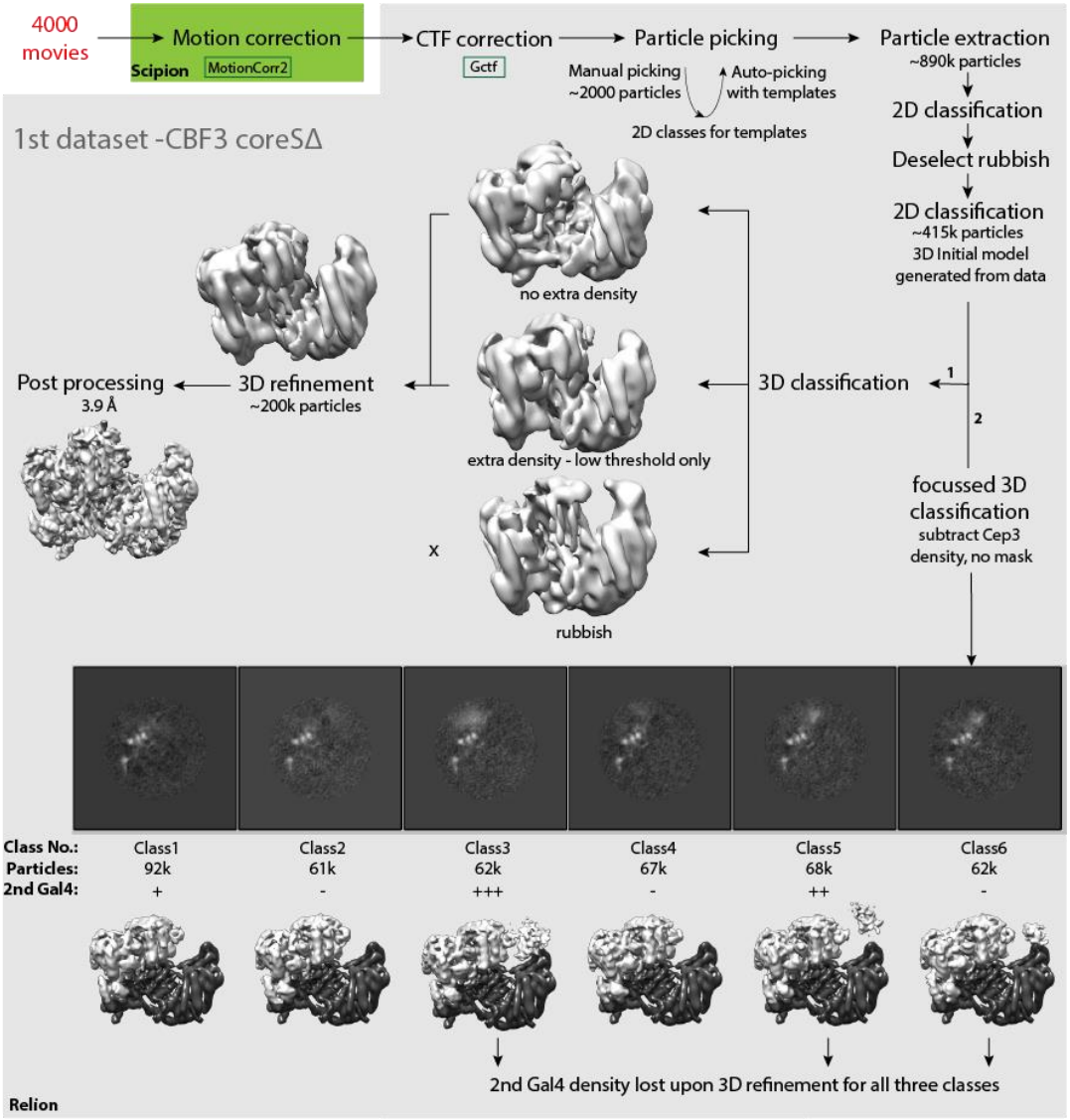

Figure 36: Processing workflow of coreSΔ dataset

An overview of the image processing steps taken while processing the coreSΔ dataset, resulting in the 3.9Å final map, as well as in some lower resolution maps showing a weak density for the second, flexible Gal4 domain.

Although the data exposed some additional diffuse density outside the core and in close proximity to the second, free Gal4-domain, this could only be observed when

using a low threshold, and was averaged out completely upon refinement (Figure 37A and B).

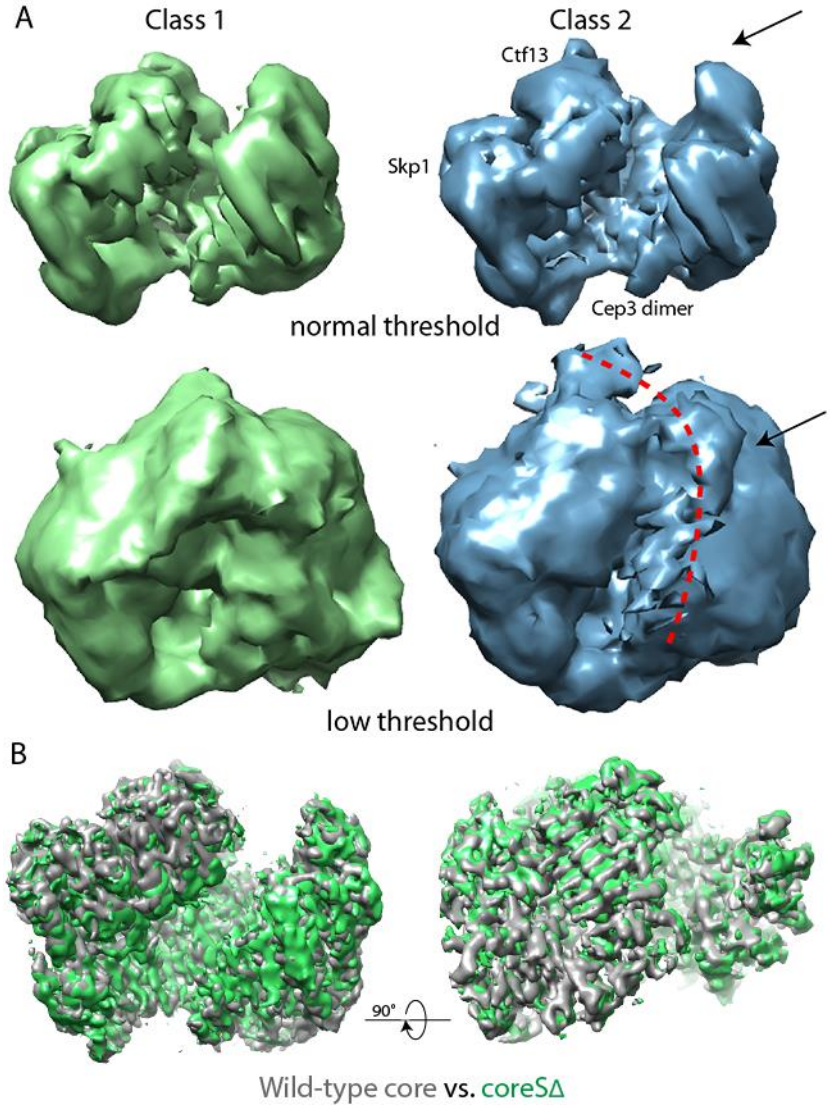

Figure 37: CryoEM density maps of coreSΔ

A Two of the three maps of the first round of 3D-classification shown at normal and low threshold. Class 1 (green) shows no extra density when viewed with a low threshold. On the other hand, class 2 (blue) shows some diffuse extra density (red dashed line) wrapping around half the structure in close proximity of the second, flexible Gal4-domain (position is indicated with the arrow). **B** Overlay of the final CBF3 core maps from the both the wild-type (grey) and SΔ (green) constructs, showing no difference of the two constructs, at least at the resolution obtained. On the left a typical side view is shown and on the right a typical top view.

Multiple rounds of 3D classification, as well as focussed 3D classification, could not improve this extra density. Nevertheless, it was possible to achieve unrefined maps showing a small extra density, which must account for the second Gal4-domain in a number of transient positions (Figure 38). Refinement of these maps, however, still lead to a loss of this density, which is inferring a high flexibility of this domain and no discrete position.

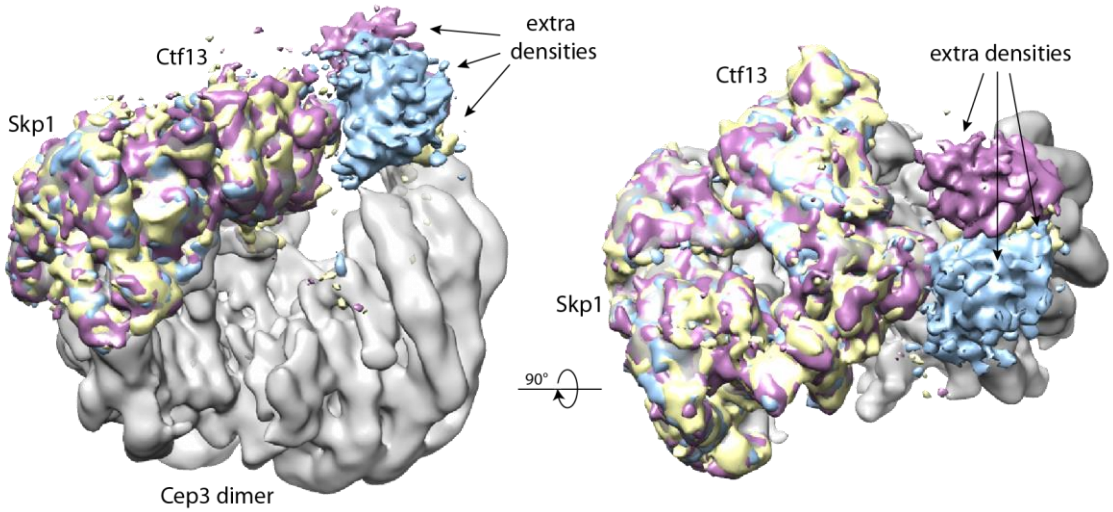

Figure 38: Focussed 3D-classification of coreSΔ

Extra density at the expected location of the second, still flexible Gal4-domain became apparent after focussed 3D classification in three classes (yellow, violet and blue). The whole core density map is shown in grey for reference. Left: typical side view. Right: typical top view.

Further 3D classification to achieve a more distinct subset, was not possible due to limiting particle number. To test if this partial stabilisation of the second Gal4-domain is because of the 37-64 Skp1 loop deletion in this construct, we attempted the same procedure of focussed 3D classification to the 2nd dataset of the wild-type core. If the second Gal4-domain is only partially stabilised in the coreSΔ structure (which is DNA binding sufficient) but not in the wild-type core structure (which cannot bind to DNA), then this stabilisation could rationalise the difference in activity. Especially, considering that the rest of the two structures are identical to each other (Figure 37B). Focussed 3D classification of the core data, however, showed similar densities of the second Gal4-domain, disproving above mentioned hypothesis. Indeed, as the initial particle number of the 2nd wild-type core complex dataset was about four times higher, more classes could be calculated, leading to a better separation of different orientations of the second Gal4-domain (Figure 39A). Refinement of some classes was therefore possible, and the best final map had an overall resolution of 3.9Å (Figure 39B). The local resolution of the second Gal4-domain, however, is poor, making the fitting of a Gal4-domain ambiguous and modelling of the atomic structure impossible (Figure 39C and D). Therefore, it is not possible to predict the likely path of DNA. Also, as both the wild-type core and coreSΔ produced similar maps, a large

conformational change to accommodate DNA binding and explain the difference in activity can be excluded.

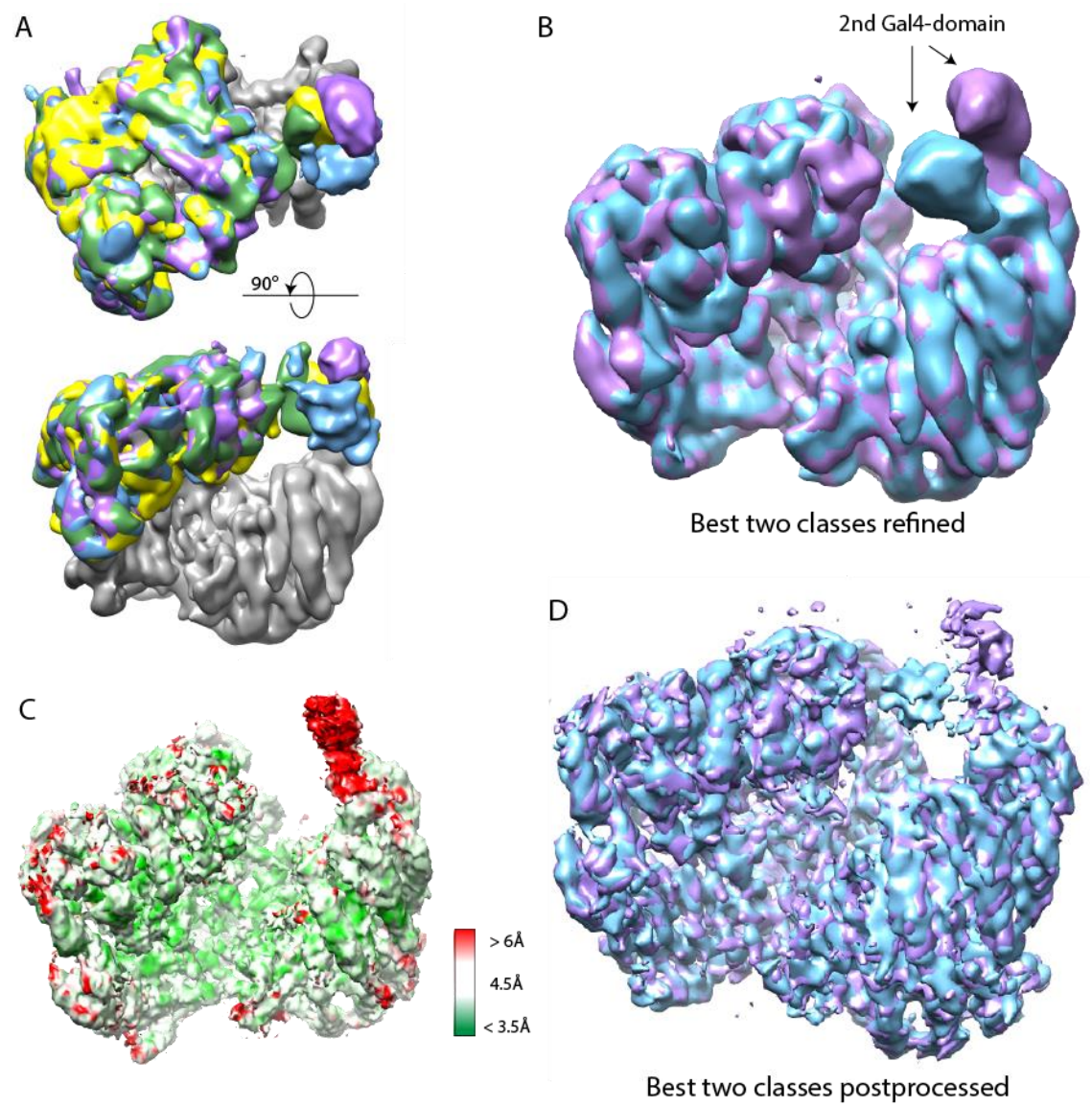

Figure 39: Focussed 3D-classification of the core wild-type data

A Four best classes showing an extra density for the 2nd Gal4-domain depicted in violet, blue, yellow and green. The core structure is shown in grey as a reference. The top shows a typical top view and the bottom a typical side view. **B** Best two classes could be refined without losing the density. **C** One of the two best classes coloured according to local resolution, showing the low local resolution at the 2nd Gal4-domain. **D** Final maps of the two best classes, showing the broken and undefined density for the 2nd Gal4-domain, making it impossible to build or fit a model and to predict the likely path of DNA.

5.2 Core complex formation

The literature describes the assembly of CBF3 as occurring step-wise. First Skp1 binds to Ctf13 to form an initial unstable sub-complex. This somehow gets activated and subsequently binds to Cep3 to form the core complex. The core is then able to interact with Ndc10, forming the full complex, which can bind to the centromere.

5.2.1 *In vitro* assembly of the core complex

To test this theory, an *in vitro* complex formation and pull-down assay was attempted. As it is impossible to express Ctf13 alone, both Skp1 and Ctf13 was co-expressed under the same conditions as the full complex, with the exception of the tag being located on the N-terminus of Skp1. Surprisingly, a good yield of Skp1/Ctf13 could be achieved and co-purified (Figure 40A), somehow speaking against the hypothesis that this initial complex gets readily degraded in a proteasome-associated manner. Cep3, was either expressed alone in budding yeast or a co-expression without Sgt1 was used, which produces an excess of Cep3. This can easily be separated via heparin and further purified (Figure 40B).

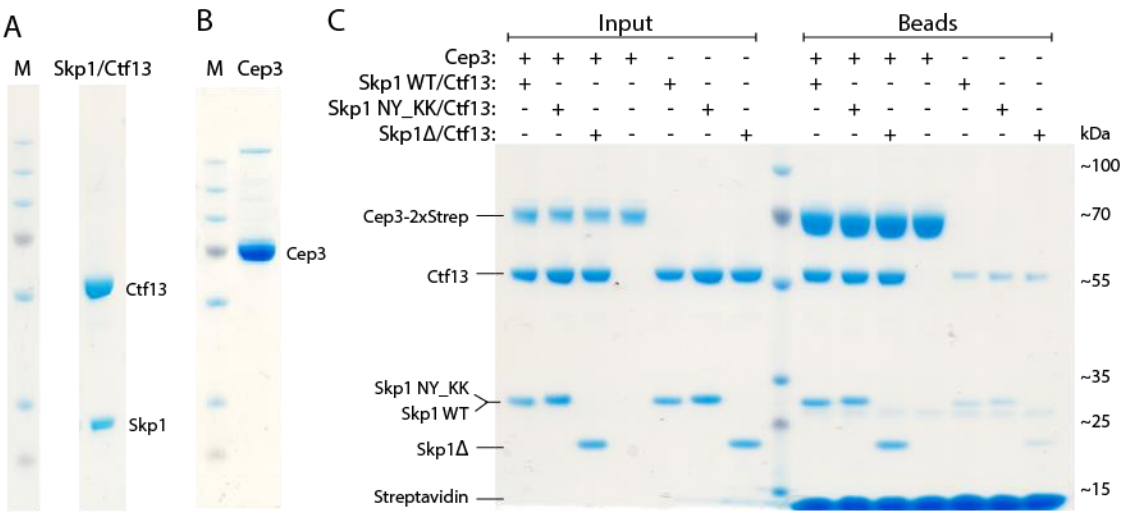

Figure 40: In vitro core assembly assay

SDS-PAGE/Coomassie of the peak fraction from size exclusion chromatography of (**A**) Skp1/Ctf13 purification and (**B**) Cep3 alone. **C** Streptavidin-beads pull-down with purified tagged Cep3 and untagged Skp1/Ctf13. Three different Skp1 constructs were tested: wild-type (Skp1 WT), residues N139 and Y140 mutated (Skp1 NY_KK) and Skp1 with the 37-64 loop deleted (Skp1Δ). The left side of the SDS-PAGE/Coomassie shows the input samples

and the right side the pull-down (Beads). The lowest band corresponds to Streptavidin, which dissociates from the beads after boiling.

Tagged Cep3 and untagged Skp1/Ctf13 could then be used for a Streptavidin beads pull-down assay, to test for core complex formation *in vitro* without any potential assembly factors present. As the material was present, not only wild-type Skp1/Ctf13 complex, but also Ctf13 in complex with Skp1Δ (37-64 loop deleted) and Skp1 NY_KK (residues N139 and Y140 mutated to two lysines) were tested. As shown in Figure 40C, all three constructs could form the core *in vitro*. This formation is therefore independent of the Skp1 37-64 loop, as well as mutations in residues N139 and Y140. Latter was unexpected given that the same mutations disrupt complex formation *in vivo* (see Figure 27D). Nevertheless, the assay proved that Cep3 and Ctf13/Skp1 form the core complex spontaneously without any additional factors needed.

5.2.2 Phosphorylation status of Skp1 F-box binding domain

As described in more detail in 4.4.3, the core structure shows a substantial difference in the conformation of the Skp1/Ctf13 F-box interaction site. It is also the only Skp1/F-box structure, which was prepared from protein expressed in an eukaryotic expression system, rather than bacteria, opening the possibility of posttranslational modifications, such as glycosylations or phosphorylations. As the latter have been indicated of importance in complex formation, as well as the fact that there are two suggested phosphorylation sites in the Skp1/F-box interaction region (beginning of Skp1 helix α7 (S162) and α8 (T177)), it was thought that phosphorylations of one or both of these sites might account for the difference in the structure and be of importance for complex formation. The local resolution in the cryoEM map, unfortunately, was too low to see any evidence of phosphorylation. Hence, Skp1 phospho-mutants (see Table 8) were co-expressed with Ctf13 and a pull-down from lysate was conducted, to see if there is a difference in the binding between Ctf13 and any of the Skp1 mutants. If phosphorylation is necessary for the interaction, then one would expect the alanine mutants (Skp1 SA, TA and SATA) to lose the ability to pull-down Ctf13. However, all alanine mutants, were able to interact and pull-down Ctf13, ruling out that a phosphorylation on any of these two sites is essential for complex formation of Skp1 and Ctf13 (Figure 41). Additionally, a double phosphomimic

mutant of Skp1 (Skp1 SETD) also interacted with Ctf13. The only mutant which showed reduced binding was Skp1 TD, whereas Skp1 SE showed no difference (Figure 41).

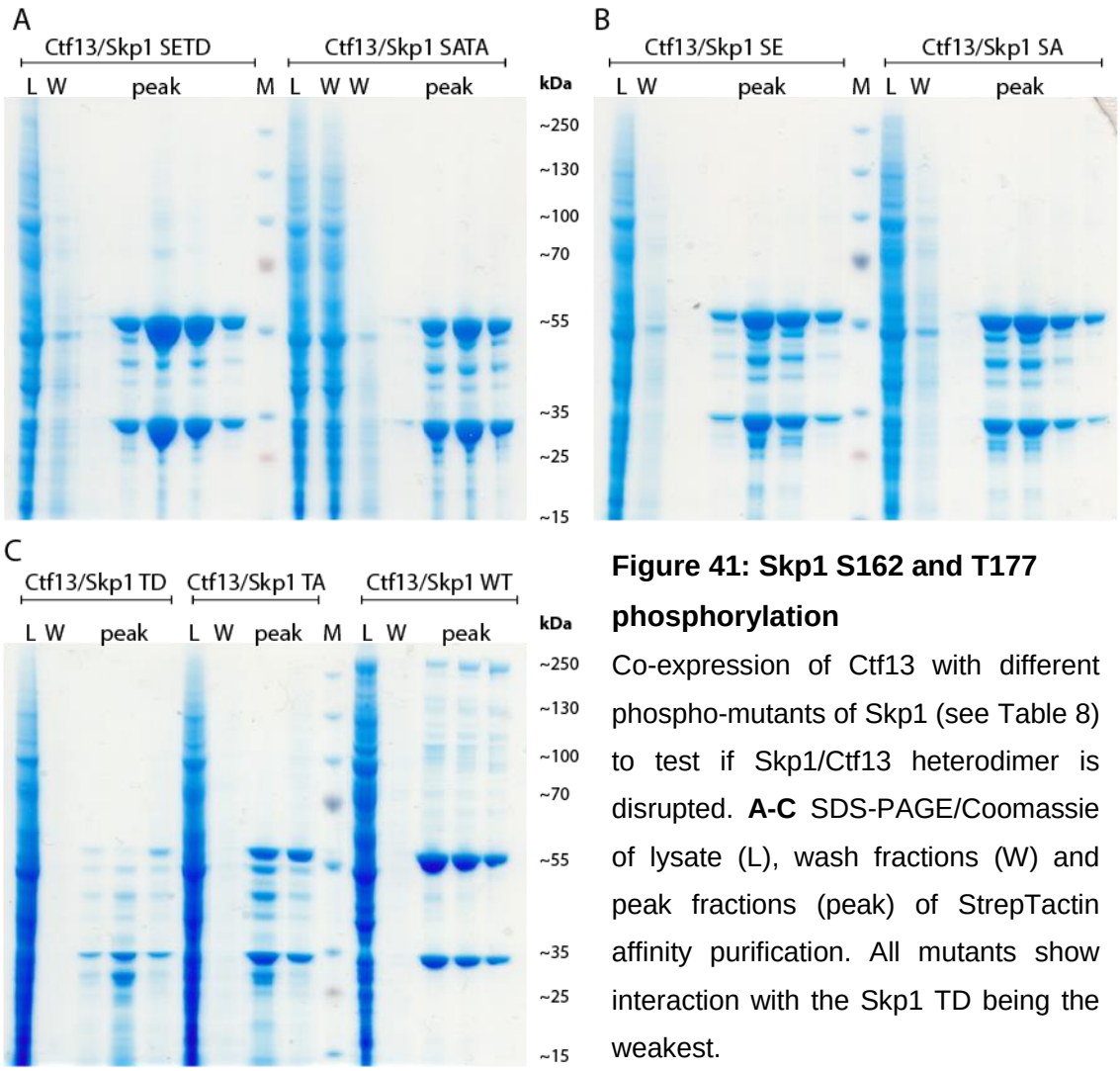

Figure 41: Skp1 S162 and T177 phosphorylation

Co-expression of Ctf13 with different phospho-mutants of Skp1 (see Table 8) to test if Skp1/Ctf13 heterodimer is disrupted. **A-C** SDS-PAGE/Coomassie of lysate (L), wash fractions (W) and peak fractions (peak) of StrepTactin affinity purification. All mutants show interaction with the Skp1 TD being the weakest.

5.2.3 Phosphorylation status of Cep3 S575

As described in 1.4.4 there is a phosphorylation site on Cep3 at residue S575, which influences DNA motility upon DNA-damage signalling (Strecker et al., 2016). Mass spectrometry analysis, carried out by the Proteomics STP, indeed showed evidence of a phosphorylation at this site. To test if this is either needed for complex formation or influences DNA-binding of the CBF3 complex phospho-mutants (Cep3 SA and SE) were tested. Co-expression of these with the other CBF3 subunits led to successful purification of the full complex, comparable to wild-type expressions,

ruling out a significant influence on complex formation (Figure 42A-C). EMSA was carried out to test a potential influence on DNA binding ability of these two constructs. However, both mutants bound to the centromeric DNA, and no difference was recorded compared to wild-type core complex (Figure 42D).

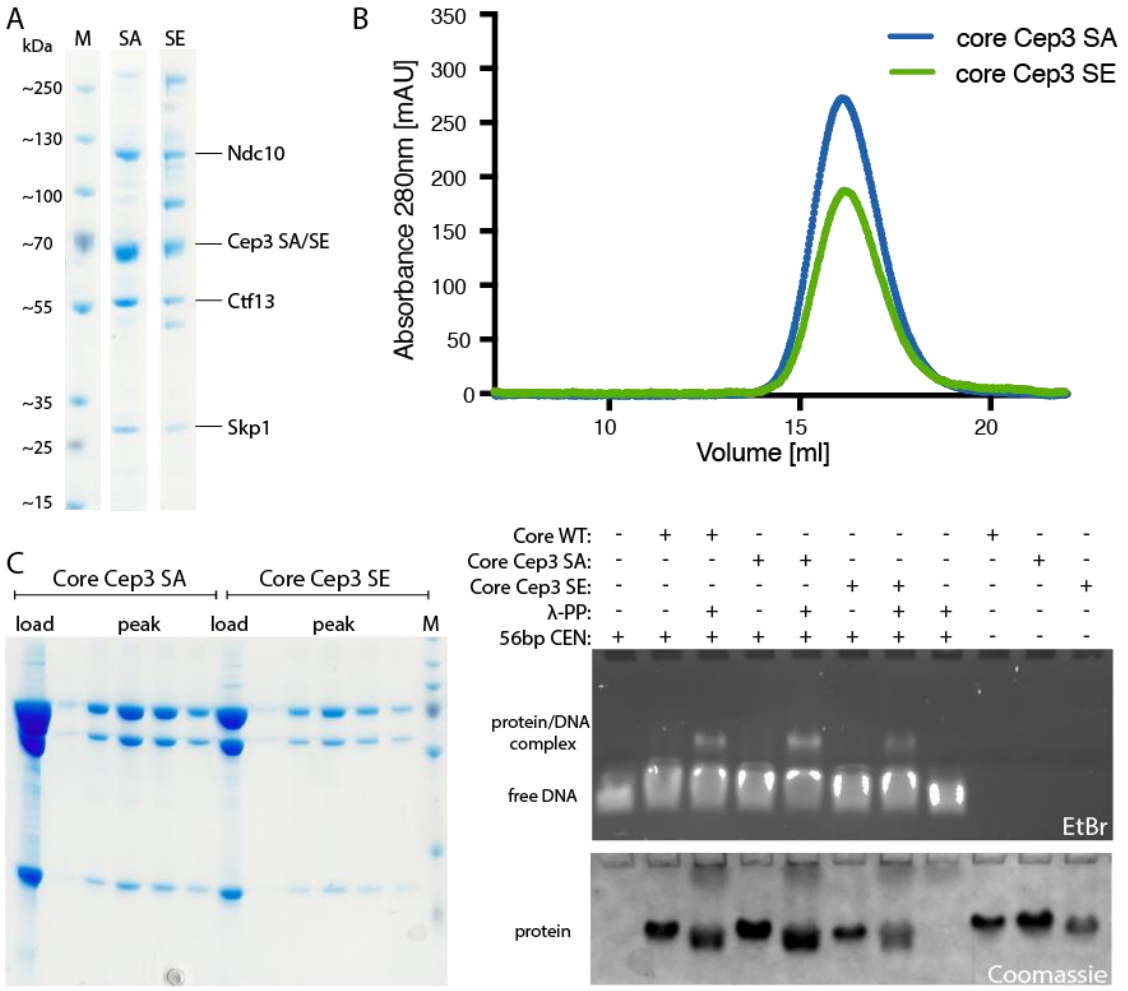

Figure 42: Analysis of core complex containing phospho-mutants of Cep3 S575

A SDS-PAGE/Coomassie stain of initial affinity purification of CBF3 containing Cep3 SA and SE phospho-mutant. **B** Size exclusion chromatography profiles of both core complexes as indicated. **C** SDS-PAGE/Coomassie stain of final size exclusion chromatography of the load and peak fractions from both core Cep3 SA and SE phospho-mutant. **D** EMSA DNA-binding assay comparing both mutants.

5.3 Interaction between CBF3 core and Ndc10

Besides many insights from the core structure, the question of how is the full CBF3 complex formed and in what way is Ndc10 interacting with the other three subunits, remained. As shown in 3.3, despite a stable enough interaction between Ndc10 and the core to withstand size exclusion chromatography, as well as a generally good elution profile, negative stain EM revealed non-optimal particles for structural elucidation. This might come from intrinsic characteristics of the Ndc10 protein, be resultant of purification strategy or might mean Ndc10 is lacking another interaction partner, to become more structured. To overcome this problem, a strategy, more typical of crystallography, was used, and various truncated Ndc10 constructs were tested, in the hope to achieve a better interaction and more suitable particles. Beside this, biochemical studies, as well as cross-linking mass spec were conducted to get a better understanding of the interaction between the largest CBF3 protein and the core.

5.3.1 Ndc10 NTD and CTD interaction with the core

The first two constructs tested were the NTD (residues 1-551), of which there is a crystal structure available, and the CTD (residues 551 to 956), which also stably expresses and purifies on its own. Both were co-expressed with the other three subunits in yeast, in the same way as the wild-type CBF3. In agreement with previous studies, which show that the N-terminal of Ndc10 is important for the interaction (Cho and Harrison, 2012), only co-expression of the NTD was successful, whilst the CTD could not be detected (Figure 43A). Even so, it was noticed that the molar ratio between the core and the NTD was not as expected (2x Ndc10, 2x Cep3, 1x Ctf13 and 1x Skp1), but a significantly smaller amount of NTD was detected. This hints on a weak interaction, or instability of this construct. Also, Ndc10 NTD lacks the dimerization domain, which could be a hindrance in its interaction to the core complex. Further purification of this complex was carried out identical to the wild-type complex and size exclusion chromatography revealed a weak interaction between the core and Ndc10 NTD (Figure 43B). Nevertheless, negative stain EM was attempted, but as expected no improvement was seen. In fact, this sample seemed to consist of core particles as well as smaller particles which could account for the NTD alone, as well as aggregates (Figure 43C).

As it was not possible to establish if the CTD of Ndc10 was successfully expressed in above described co-expression and co-purification experiment, a pull-down assay with purified protein was done to confirm this result. Both NTD and CTD were expressed in *E.coli*, and as a control also full length Ndc10 was purified from the co-expressed full complex via the Heparin step (see 2.5.2 for details, note that a small amount of core will still be present in this sample). As expected, only Ndc10 NTD showed an interaction with the core, whereas the CTD did not. Full length Ndc10 also interacted with the core, however, nearly the same amount did also interact with the streptavidin beads in the control reaction (Figure 43D). Nevertheless, the assay clearly showed that Ndc10 NTD can interact with the core, whereas the CTD alone cannot.

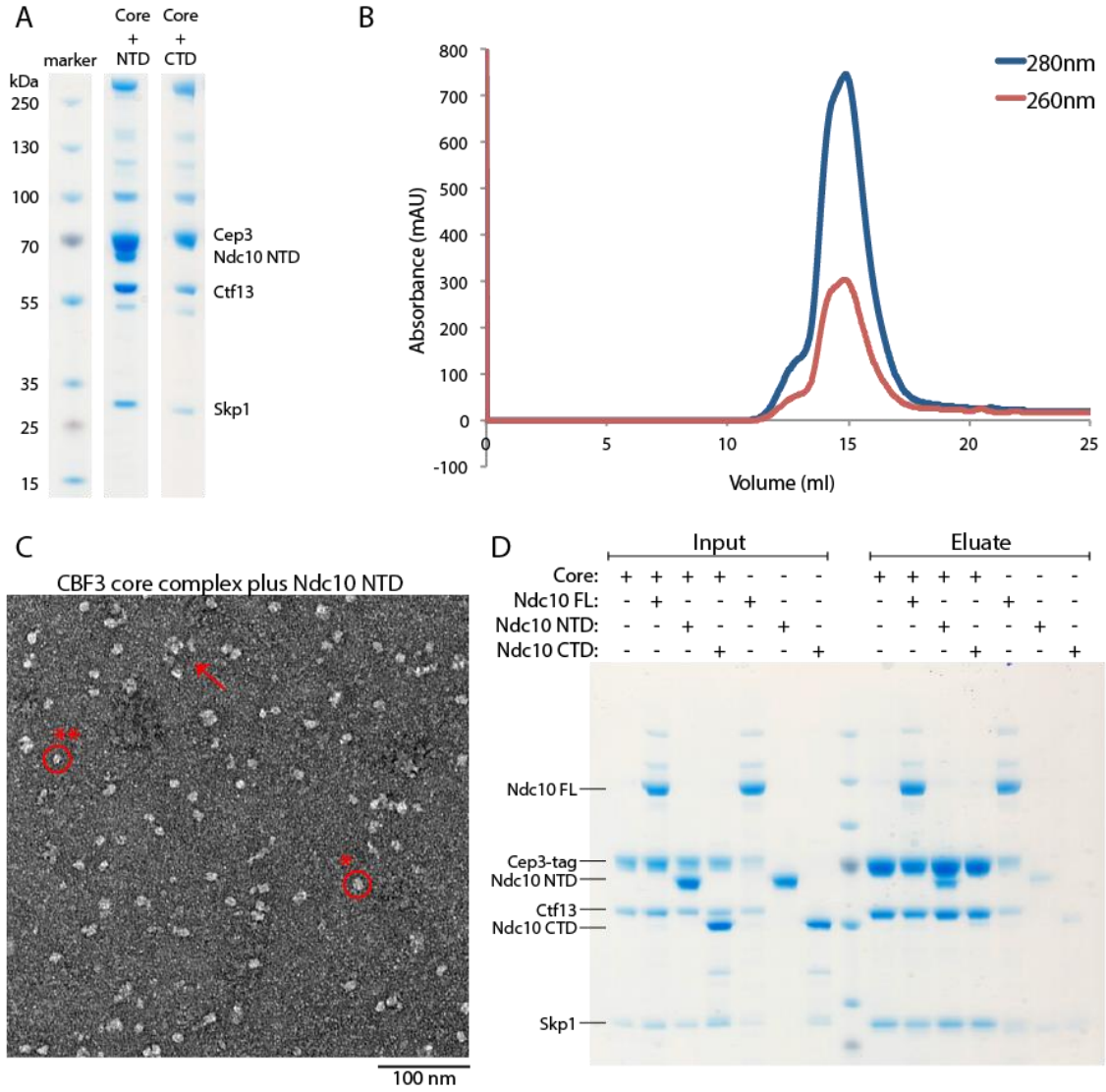

Figure 43: Interaction between core complex and Ndc10 NTD/CTD

A SDS-PAGE/Coomassie stain of the elution after affinity purification of co-expression of tagged core with untagged Ndc10 NTD (residues 1-551; middle panel) or CTD (residues 551-956; right panel). The CTD would be ~47.3kDa and does not co-purify with the core. The NTD is ~64.6kDa and co-elutes with the core, however not in an equimolar ratio. **B** Suboptimal size exclusion profile of core plus Ndc10 NTD, exhibiting a shoulder and two joined peaks. **C** Negative stain EM micrograph of core plus Ndc10 NTD, made straight after size exclusion chromatography. The red circles highlight a typical core particle (one asterisk) and a smaller particle probably accounting for the NTD (two asterisks). The arrow shows some aggregation. **D** SDS-PAGE/Coomassie of a pull-down assay with tagged core and Ndc10 full-length (FL), NTD and CTD. As shown in **A** the CTD does not interact with the core, whereas the NTD does. Full-length Ndc10 also sticks to the beads unspecifically.

5.3.2 Other Ndc10 constructs

As the NTD of Ndc10 is essential but not sufficient for the binding to the core complex, limited proteolysis was conducted to find a more stable construct. Both Ndc10 alone and in complex with the other three CBF3 proteins, as well as the core complex without Ndc10 was tested. Two stable degradation bands could be detected for Ndc10. Band 1 was more evident in the sample including the other three subunits, whereas the more stable band 2 appeared in both (Figure 44A). Besides Ndc10, also Ctf13 is degraded very quickly and entirely, whereas Cep3 and Skp1 seem to be more stable, even after 15min of digest. Furthermore, there is no difference in the degradation pattern of Ctf13 with and without Ndc10 present. This argues against a tight and vast binding interface between Ctf13 and Ndc10, which would lead to some protection against degradation of Ctf13.

The two Ndc10 degradation bands were tested with mass spectrometry (analysis done by the Proteomics STP) and it was found that both consist of a C-terminal truncation of Ndc10. Three constructs, ending at residues 848, 777 and 683, were designed accordingly and co-expression was tested. All three of these constructs co-purified with the core (Figure 44B), however, all exhibited a similar pattern as the full length Ndc10: a fair amount was seen in the wash of the affinity column and the yield obtained was small. Their behaviour during purification was similar as well, except that of the smallest construct, which was unstable and nearly all was lost consequently (Figure 44C). Negative stain analysis again, showed no improvement of either of these constructs compared to the full-length protein (Figure 44D).

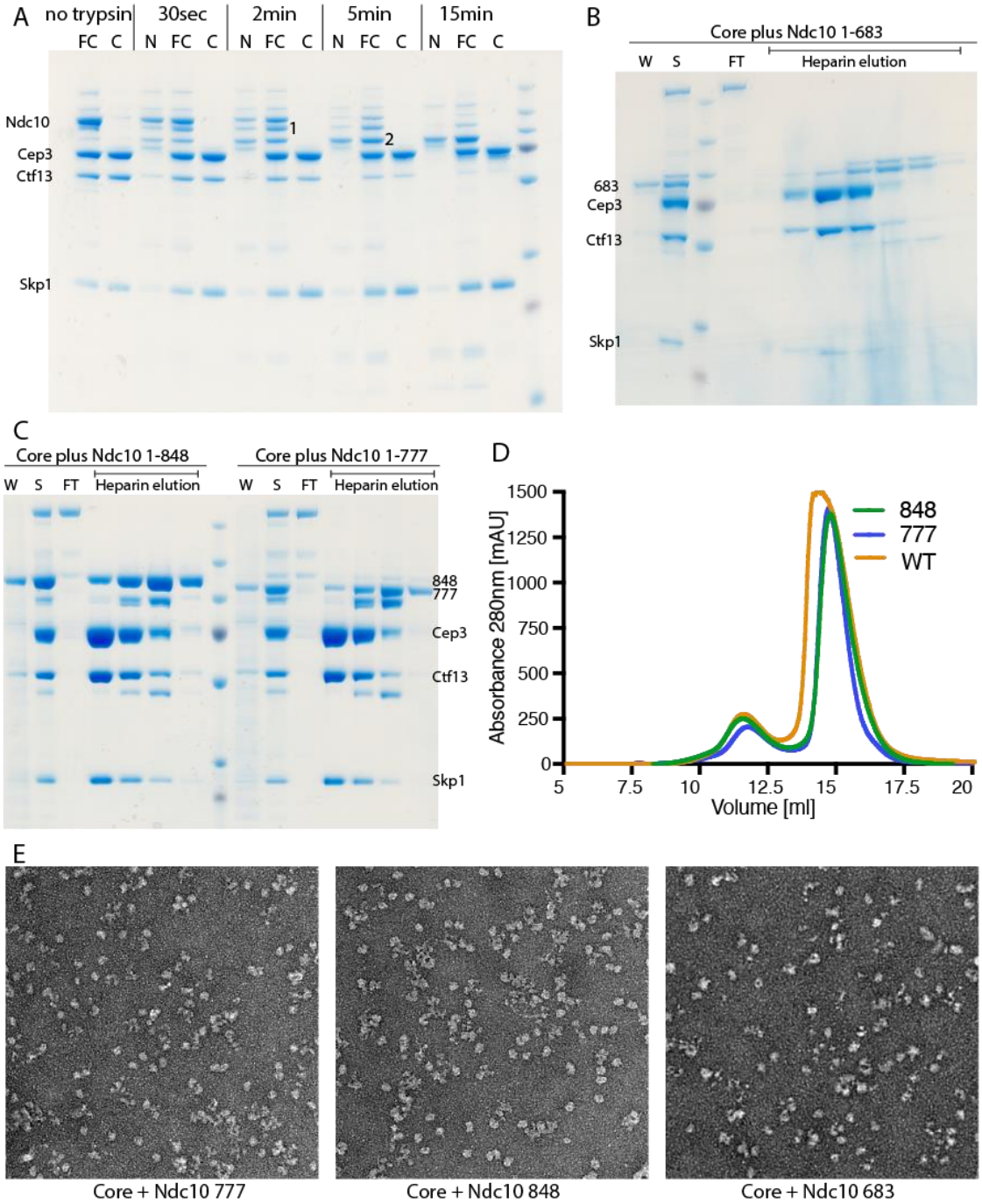

Figure 44: Various Ndc10 constructs and their interaction with the core

A SDS-PAGE/Coomassie stain from a limited proteolysis experiment, exhibiting two stable Ndc10 bands (1 and 2). SDS-PAGE/Coomassie stain of the initial purification steps of three Ndc10 constructs: 1-683 (**B**), 1-848 and 1-777 (**C**). The first lane corresponds to the wash (W) fraction of the StrepTactin column, showing that all three constructs partly dissociate from the stably bound core; the second lane is the peak fraction of the StrepTactin elution (S), followed by the typical heparin elution peak. **D** Size exclusion profiles of the wild-type

(WT) core compared to two of Ndc10 constructs (1-777 and 1-848). **E** Negative stain micrographs of all three constructs, not showing any improvement of particles.

5.3.3 Difference of Ndc10 interaction stability with core and coreSΔ

During the purification of coreSΔ for DNA binding, as well as structural studies (see 5.1.5), it was noticed that every purification consistently yielded more Ndc10 than wild-type CBF3 (Figure 45A).

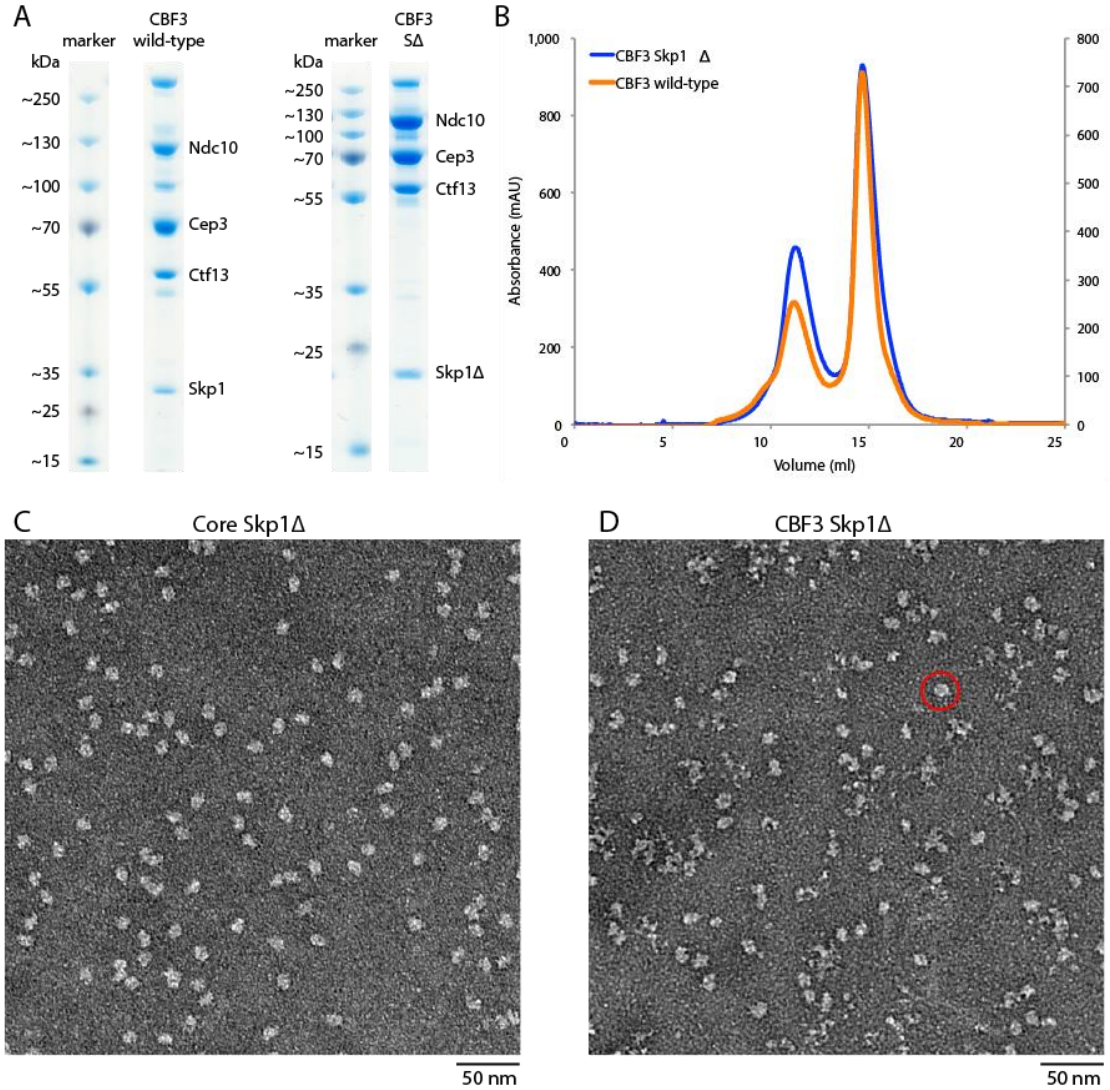

Figure 45: Difference between wild-type CBF3 and CBF3 SΔ

A SDS-PAGE/Coomassie stain of the initial purification of both, showing an increased Ndc10 yield for the CBF3 SΔ co-expression and purification. **B** Size-exclusion chromatography profiles showing no difference other than more CBF3 SΔ full complex. Negative stained micrographs of core (**C**) and full CBF3SΔ (**D**). Whereas the core particles look identical to

wild type, the full complex shows no improvement and one can distinguish between core particles (red circle) and additional diffuse density.

This lead to the hypothesis that Ndc10 might bind better to coreSΔ than core complex with the phosphorylated Skp1 37-64 loop. This could also mean that CBF3 SΔ would produce better particles, suitable for structural studies. To test this, CBF3SΔ was subjected to size-exclusion chromatography and negative stain EM. Size exclusion profile, however, showed no improvement compared to the wild-type CBF3, and consistently, negative stained particles were identical (Figure 45B-D). A hypothesis of how coreSΔ binds more Ndc10 is proposed in the discussion (see 7.1).

5.3.4 Cross-linking mass spectrometry

As structural studies of the full CBF3 complex were not feasible, but structural data of the core and part of Ndc10 is available, cross-linking in combination with mass spectrometry was attempted. This method can give indications of protein/protein interactions, by identifying amino acid pairs which are in close proximity to each other. This data can then be used, to model the interaction, especially if additional structural, or biochemical data is available to validate the crosslinks.
CBF3 and core protein was expressed and purified as described in 2.5.2. Straight after size exclusion chromatography the protein was crosslinked with disuccinimidyl sulfoxide (DSSO) and subjected to mass spectrometry analysis. All sample preparation after the crosslinking, mass spectrometry and data analysis were carried out by the Proteomics STP. 19 and 35 inter- and intra-molecular crosslinks were found in the core and CBF3 sample respectively (Figure 46A and B; Table 11). Both samples had many overlapping crosslinks between and within the core subunits. All self-links, situated within the core or Ndc10 NTD crystal structure are feasible, as the two crosslinked peptides are in close proximity to each other. As there is no sequence assigned full model available for Ctf13, crosslinks within this subunit cannot be genuinely validated. However, all of them are within the LRRs and many in close proximity of each other and therefore, too seem viable. Self-links within Cep3 are concentrated within the Gal4-domains, as well as some between the two monomers, whereas none were found for Skp1. Crosslinks between the subunits of the core also fit very well with the cryoEM structure, as the N-terminus of Skp1, which lies in close proximity to the F-box distal end of the LRRs, shows corresponding

crosslinks. Furthermore, the Gal4-domain of Cep3 crosslinks extensively to the LRRs of Ctf13, which can be explained as the second, still flexible Gal4-domain is located in this area (see 4.4). Although there are many self-links within Ndc10, only four crosslinks could be recorded between Ndc10 and the core. Three of these were to Ctf13 and one to Cep3. Interestingly, all crosslinks to Ctf13 were from the NTD of Ndc10, in accordance with the biochemical data showing the NTD being essential for the interaction. The other crosslink was between the very C-terminus of Ndc10 and the Gal4-domain of Cep3. These results could explain why the NTD is important but not sufficient for the interaction. The possible meaning of the crosslinks between Ndc10 and the core subunits is discussed in 7.1.

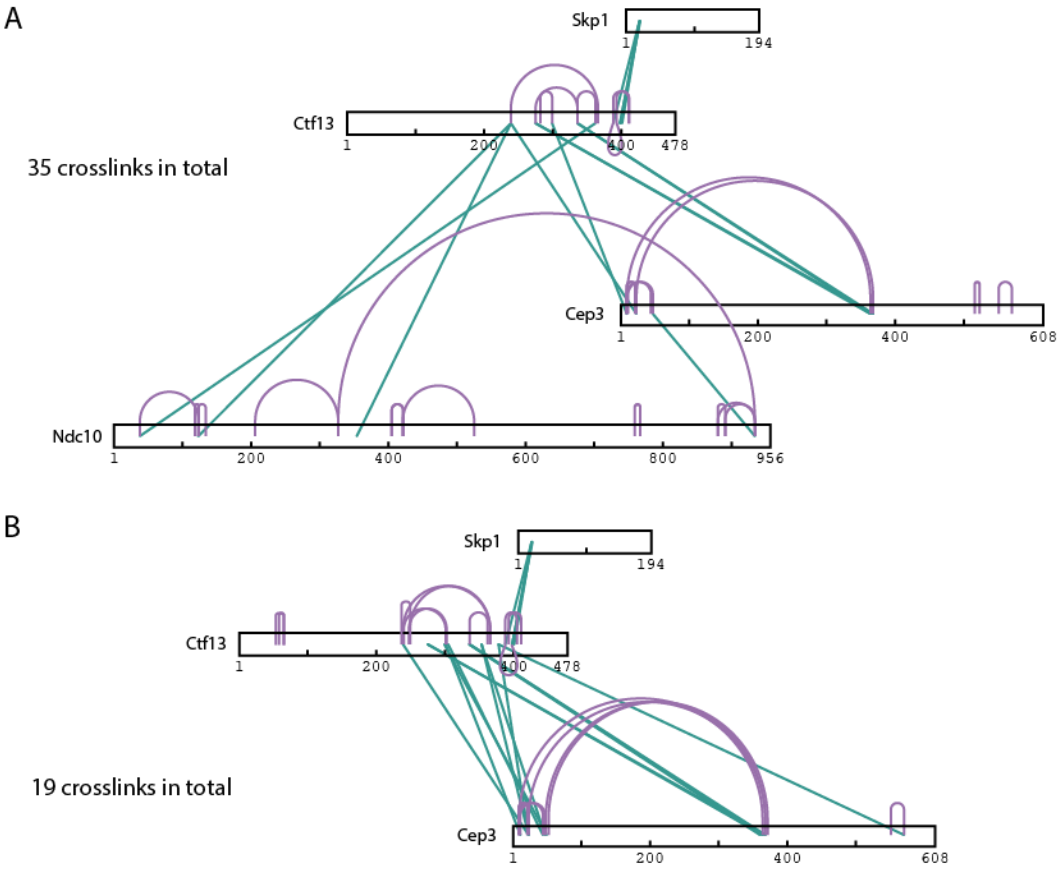

Figure 46: Cross liking- mass spectrometry

Data from the whole CBF3 (**A**) and core complex (**B**). Crosslinks are shown as lines and colour coded according to type: crosslink between subunits (green), inter- or intramolecular self-link (violet) or homo-multimeric self-link (red). Overall more crosslinks could be recorded between the core subunits when Ndc10 was not present (B), however, this might be an experimental artefact, as the whole complex tends to aggregate especially in the high concentration needed for the crosslinking reaction.

Protein 1	Protein 2	Peptide 1	Peptide 2	Seen with CBF3	Seen with core	Score (CBF3)	Score (core)
Ndc10	Ctf13	30-FQELVDSLKPR	358-SLTMKR	☐		22.86	
Ndc10	Ctf13	119-IATLKK	236-SASKR	☐		24.64	
Ndc10	Ctf13	348-YCHLPYKR	236-SASKR	☐		49.31	
Ndc10	Cep3	932-LAKFLR	45-STKLITASSSK	☐		13.84	
Ndc10	Ndc10	14-IMDVQQQQEAMSSEDRFQELVDSLKPR	124-KIIGSLNFLSK	☐		149.06	
Ndc10	Ndc10	103-REETEDLDEEEENSFKIATLK	124-KIIGSLNFLSK	☐		20.46	
Ndc10	Ndc10	119-IATLKK	125-IIGSLNFLSKLCK	☐		24.3	
Ndc10	Ndc10	204-FFKTCSEK	326-GKSLTTYPR	☐		84.57	
Ndc10	Ndc10	405-KIMNFK	422-KDPFPPPK or 411-SPYTSYSTNAKKDPFPPPK	☐		36.47 or 29.32	
Ndc10	Ndc10	755-ILEEKVTR	763-IIEQKLGSHTGK	☐		29.81	
Ndc10	Ndc10	877-SGNKSWR	884-ANCEALYKER	☐		54.44	
Ndc10	Ndc10	877-SGNKSWR	932-LAKFLR	☐		59.82	
Ndc10	Ndc10	884-ANCEALYKER	932-LAKFLR	☐		100.09	
Ndc10	Ndc10	326-GKSLTTYPR	932-LAKFLR	☐		17.1	
Ndc10	Ndc10	411-SPYTSYSTNAKK	523-IYKNK	☐		15.5	
Skp1	Ctf13	16-FTVDKK	386-TFKSLYIKEDDSK	☐		18.78	
Skp1	Ctf13	16-FTVDKK	389-SLYIKEDDSKFTK or 399-FTKFFNLR	☐	☐	37.38 or 81.88	42.46 or 22.86
Skp1	Ctf13	21-KIAER	394-EDDSKFTK or 399-FTKFFNLR	☐	☐	20.28 (first)	54.54 or 62
Cep3	Ctf13	5-TTQLKSK	295-DNPGKTISYIVK	☐		33.32	
Cep3	Ctf13	22-VKCDR	236-SASKR	☐		11.47	
Cep3	Ctf13	22-VKCDR	379-ELKVDK		☐		26.6
Cep3	Ctf13	37-GQDSECMKSTK	236-SASKR		☐		40.1
Cep3	Ctf13	37-GQDSECMKSTK	342-VAYIDLNSIVFPKNFK		☐		20.9
Cep3	Ctf13	362-SSKPPLKTLDAIR or 362-SSKPPLKTLDAIR	275-KITVR	☐		64.24 or 13.05	68.23 (first)
Cep3	Ctf13	362-SSKPPLKTLDAIR or 362-SSKPPLKTLDAIR	331-WDNLQKLVLSR	☐	☐	77.06 03 117.17	
Cep3	Ctf13	561-GLGSLVPLNKLR	379-ELKVDK		☐		30.25
Cep3	Cep3	5-TTQLKSK	21-KVKCDR or 22-VKCDR	☐	☐	12.98 or 65.83	45.93 or 73.97
Cep3	Cep3	5-TTQLKSK	37-GQDSECMKSTK or 45-STKLITASSSK	☐	☐	15.69 (first)	31 (second)
Cep3	Cep3	5-TTQLKSK	362-SSKPPLKTLDAIR	☐	☐	52.59	54.44
Cep3	Cep3	10-SKHPCSVCTR	45-STKLITASSSK	☐	☐	14.16	40.11
Cep3	Cep3	22-VKCDR	362-SSKPPLKTLDAIR	☐	☐	79.14	23.56
Cep3	Cep3	37-GQDSECMKSTK	354-DLDQYLNKSSKPPLK or 362-SSKPPLK		☐		20.45 or 17.86
Cep3	Cep3	45-STKLITASSSK	362-SSKPPLK		☐		37.48
Cep3	Cep3	514-LSKLK	519-LLWDKVQLLDSGDSFYHPVFK	☐		31.92	
Cep3	Cep3	547-IIELKNDEMFSLIK	561-GLGSLVPLNKLR	☐		80.18	
Ctf13	Ctf13	212-DVDVSGANSDENSSPSSTIKNK	236-SASKR		☐		25.96
Ctf13	Ctf13	212-DVDVSGANSDENSSPSSTIKNK	363-RVSKIK		☐		22.15
Ctf13	Ctf13	236-SASKR	364-VSKIK	☐	☐	35.62	40.44
Ctf13	Ctf13	275-KITVR	331-WDNLQKLVLSR	☐		35.62	
Ctf13	Ctf13	280-GEKLYELLINFHGFR	295-DNPGKTISYIVK	☐		15.09	
Ctf13	Ctf13	331-WDNLQKLVLSR	358-SLTMKR	☐	☐	66.66	57.99
Ctf13	Ctf13	386-TFKSLYIKEDDSK	389-SLYIKEDDSKFTK	☐		8.49	
Ctf13	Ctf13	386-TFKSLYIKEDDSK or 386-TFKSLYIKEDDSK	410-IKELDKSEINQITYLR	☐	☐	55.39 or 10.83	82.5 (first)

Table 11: CBF3 crosslinks

Detailed list of all crosslinks found between the core and full complex.

5.3.5 Ncd10 peptide array

In addition to co-expression of different Ndc10 constructs, as well as pull-down assays and crosslinking MS, a peptide array of Ndc10 was tested against tagged core complex. Fourteen areas of interaction were found, whereas the control showed no signal (Table 12 and Figure 47A and B). Of these possible interaction sites, ten were in the NTD (residue 1-550), and the other four in the CTD. Of the latter, three were close to the NTD, between residues 574 and 620, and only one at the very C-terminus at residues 924 to 930 (Figure 47C). These results correspond well with the finding that the NTD is essential for interaction, but not sufficient.

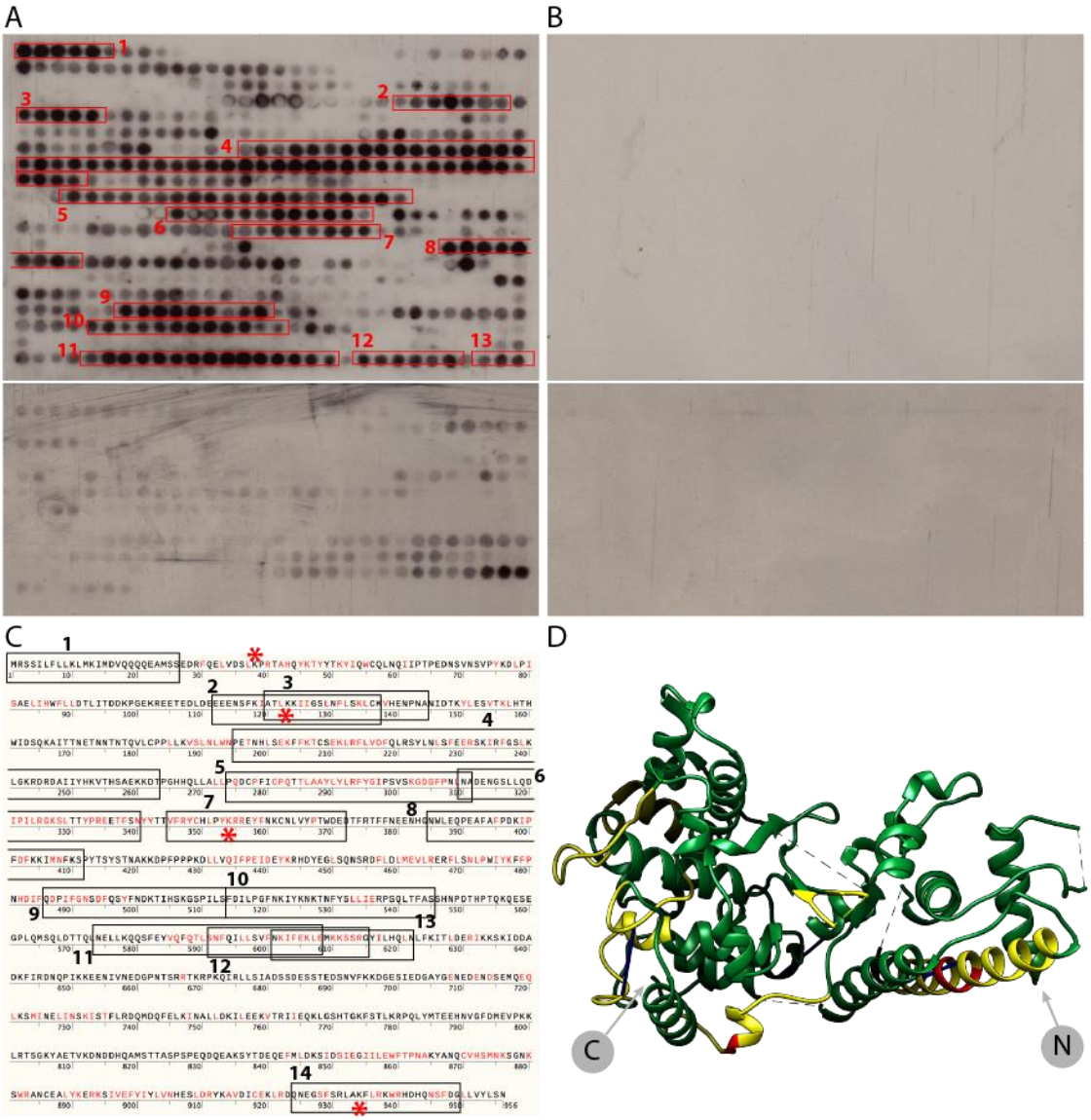

Figure 47: Ndc10 peptide array probed against with core complex

A Membrane 1 (top) and membrane 2 (bottom), developed on film using a primary Anti-Strep and secondary Anti-mouse HRP-conjugated antibody. **B** Duplicate control membranes treated equally, except for incubation without core complex. **C** Protein sequence of Ndc10, highlighting the potential areas found in the peptide array (boxes), as well as sequence conservation (red letters: fully or partially conserved). The asterisks show the four crosslinks found between Ndc10 and the core complex. **D** Crystal structure of Ndc10 NTD (4ACO), showing solvent exposed areas found in the peptide array (yellow), as well as two of the four crosslinked residues (red).

Furthermore, three of the four crosslinks between Ndc10 and either Ctf13 or Cep3 correspond with an element found in the peptide array (Figure 47C). From those areas within the crystal structure, most are entirely or partly solvent exposed, and seem to be concentrated on two sides of the structure (Figure 47D). However, it should be noted that the crystal structure only compromises the N-terminal domain forming a monomer (construct used to crystallise: residues 1-551; residues resolved in the crystal structure: residues 44-537 with some internal loops missing), whereas the full-length Ndc10 is nearly double the size and forms a homodimer. Therefore, some or all of these areas could be inaccessible in the context of full-length protein.

Table 12: Summary of peptide array data

Conservation is indicated on a scale from not conserved at all (-) to somewhat (+), partly (++) and very (+++) conserved. It is specified in which domain the areas belong, and for those situated within the published crystal structure (4ACO) it is indicated if and to what extend they are on the outside of the structure (i.e. solvent exposed).

No.	Residues	Conservation	Domain	Solvent exposed
1	1 to 26	-	NTD	n.a.
2	113 to 137	+	NTD	yes
3	121 to 144	+	NTD	yes
4	195 to 263	++	NTD	partly yes
5	274 to 312	+++	NTD	partly yes
6	310 to 340	++	NTD	partly yes
7	345 to 371	++	NTD	yes
8	386 to 412	+	NTD	no
9	486 to 514	+	NTD	yes
10	514 to 545	+	NTD	n.a.
11	574 to 608	+	CTD (centralCTD)	n.a.
12	591 to 615	+++	CTD (centralCTD)	n.a.
13	599 to 620	++	CTD (centralCTD)	n.a.
14	924 to 949	+	CTD (CTDshort)	n.a.

5.4 Interaction between CBF3 and the Cse4-nucleosome

Both the centromeric nucleosome and CBF3 have been shown to bind to the centromere in yeast. How both of these rather large complexes can simultaneously interact with the short centromeric sequence is a matter of some debate.

Here, it was tested if CBF3 can recognise and bind to centromeric DNA wrapped around a histone octamer *in vitro*, as has been shown for the CCAN protein Mif2 (Xiao et al., 2017). Therefore, nucleosomes were reconstituted both with centromeric DNA and Widom 601 DNA (Table 9), as well as with the canonical H3 and centromere-specific Cse4 full-length and a construct lacking the divergent N-terminal tail (Cse4 HFD). EMSAs were carried out to see if addition of CBF3 results in a shift, as expected if there is an interaction. First nucleosome formation was tested with a native agarose gel (Figure 48A). All three different histone octamers formed nucleosomes with Widom 601 DNA, however, only H3 containing octamer has formed a distinct nucleosome on centromeric DNA (147bp or 185bp). Cse4 HFD containing octamer seemed to have formed an inconsistent protein DNA complex, smearing across the gel, and Cse4 full-length did not form a nucleosome with the centromeric DNA at all. These results might be expected to some extent, as centromeric DNA is very A/T-rich and therefore suboptimal for nucleosome formation, and so is the Cse4 full-length protein with its relatively long N-terminal tail. Nevertheless, two EMSA experiments were made. The first one tested for binding of CBF3 with either H3 nucleosome on Widom 601 DNA or a chimeric H3 nucleosome on centromeric DNA (Figure 48B). If CBF3 and centromeric nucleosome simply bind the same stretch of DNA but have no direct interaction between them, then CBF3 should be able to bind to the chimeric but not the Widom 601 H3 nucleosome. However, no binding of both nucleosomes was detected. In fact, CBF3 preferentially bound to the free DNA present in the nucleosome sample, independently from DNA sequence. The second experiment was designed to see if CBF3 interacts with any of the three different nucleosomes on Widom 601 DNA (Figure 48C). In this case centromere-specific DNA sequence is not present, however, CBF3 also non-specifically interacts with DNA and might show a binding to the centromere-specific Cse4 independently to DNA sequence. If that is true, one would expect a shift of CBF3 mixed with Cse4 full-length and/or Cse4 HFD but not with H3. However, no clear shift was detected in any of the three cases, and as before CBF3 seems to

prefer to bind to the free DNA present in the nucleosome sample. Nevertheless, by careful examination, CBF3 mixed with H3 nucleosome has bound nearly all of the free DNA, whereas in both Cse4 nucleosomes, some free DNA is left. One should note that, there is approximately the same amount of free DNA present in all three samples. Similarly, comparing the Coomassie stain intensity between the nucleosomes bands alone, and mixed with CBF3 full complex, both H3 and Cse4 HFD stay the same, but there is a reduction in intensity of the nucleosome band with the Cse4 full-length sample (Figure 48C). Input samples on a SDS-PAGE gel are shown in Figure 48D.

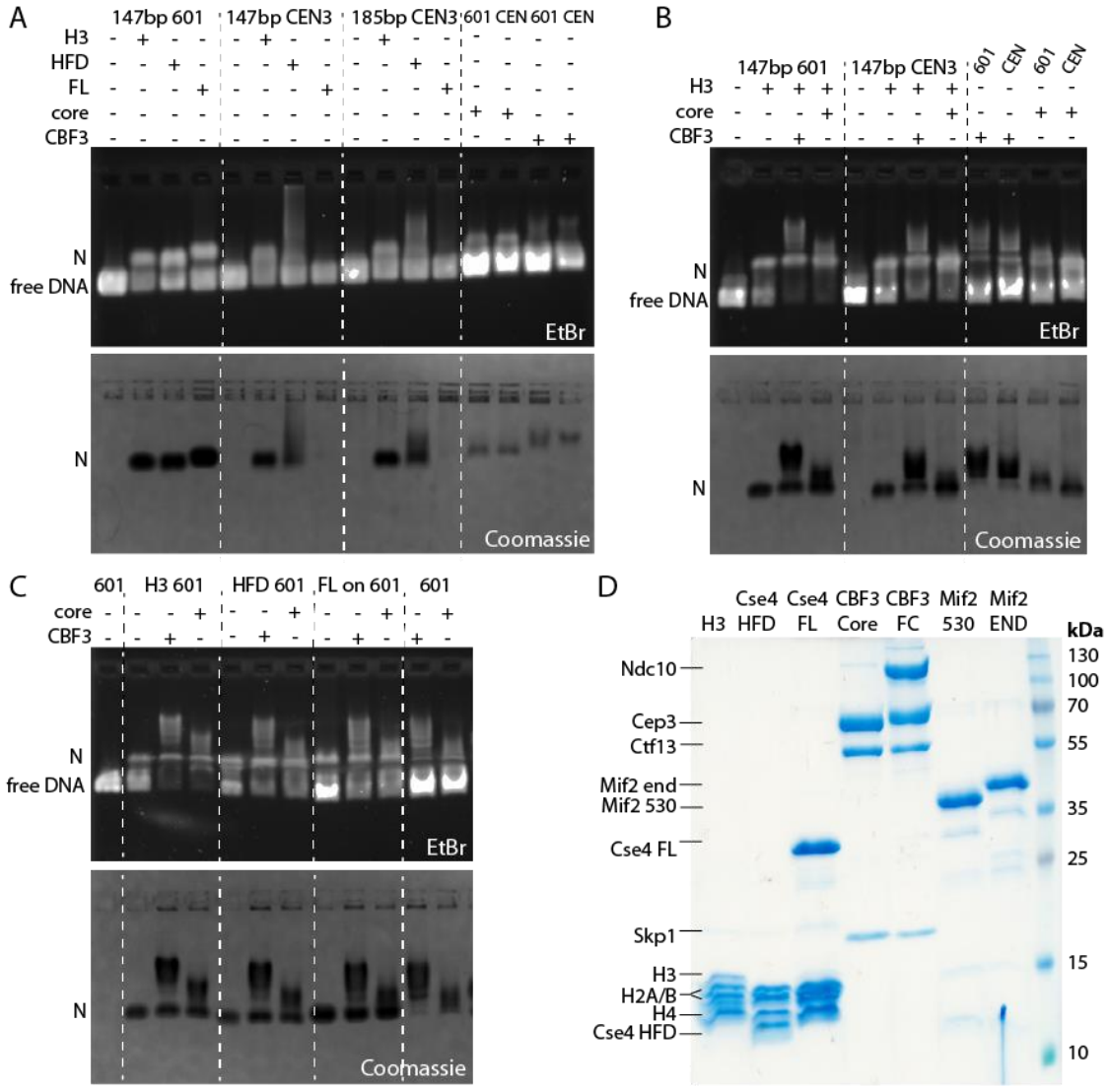

Figure 48: EMSA studies of CBF3/nucleosome interaction

A Formation of different nucleosomes: a combination of H3 octamer (H3, H4, H2A, H2B), HFD (Cse4 HFD, H4, H2A, H2B) or FL (Cse4 FL, H4, H2A, H2B) with either 147bp Widom

601, 147bp or 185bp CEN3 DNA were prepared and nucleosome formation was tested with a native agarose gel. Whereas all histone variants formed nucleosomes with 601 DNA, only H3 formed a distinct nucleosome on both centromeric DNAs. **B** EMSA assay to test if core or full CBF3 complex shows interaction with a canonical H3 nucleosomes on 601 DNA or a chimeric H3 nucleosome on CEN3 DNA. As controls, free DNA (lane1 and 5) and core/CBF3 mixed with free DNA (lanes 9-12) is run. **C** EMSA with three types of nucleosomes, H3, Cse4HFD (HFD) and Cse4 full-length (FL) on 601 DNA, and core and full complex (CBF3) showing no interaction. As in B core and CBF3 preferably bind to the free DNA present rather than any nucleosome. As controls, free DNA (lane 1) and core and CBF3 mixed with DNA (lanes 11 and 12). **D** SDS-PAGE/Coomassie stain of all input samples used in A-C, as well as two Mif2 constructs used in Figure 47.

As a control of principle, the assay was repeated with Mif2, as published before (Xiao et al., 2017). Two Mif2 constructs were tested, Mif2 530 and Mif2 END, which also included the last 19 C-terminal residues. In this case a clear shift of the nucleosome band can be seen when adding Mif2 530 or Mif2 END (Figure 49A and B). This shift becomes apparent already with the lowest concentration of Mif2, and in accordance with the published assays, Mif2 seems to interact better with Cse4 nucleosomes than the canonical H3 nucleosome. Hence, the assay is working but CBF3 shows no or very little binding to the nucleosome, however, further improved analysis will be necessary to confirm these preliminary tests and these are discussed in 7.3.

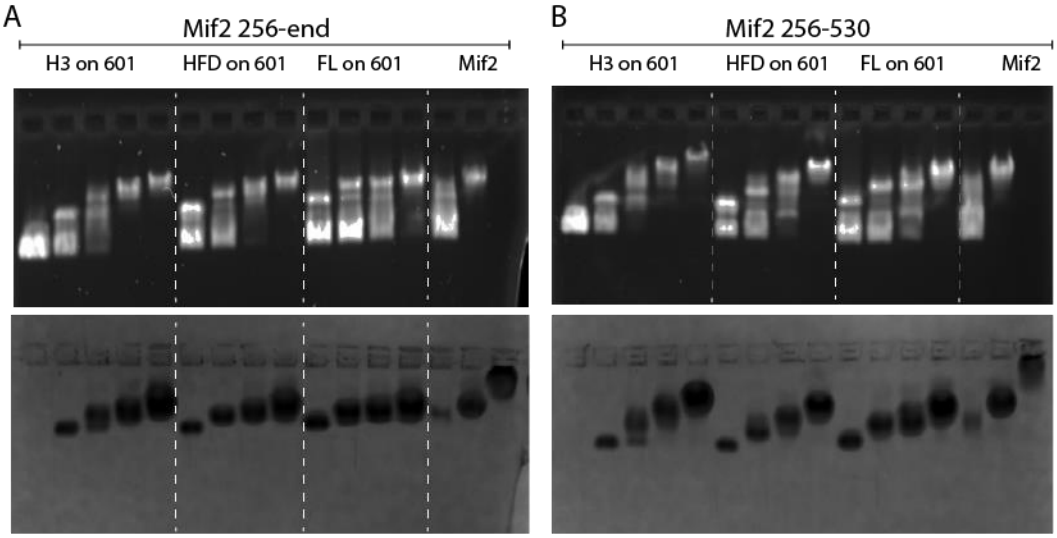

Figure 49: Mif2/nucleosome interaction studies

EMSA assays as proof of principle with the same nucleosomes used for Figure 46 and two different Mif2 constructs: 256-end (**A**) and 256-530 (**B**). The latter was used by Xiao et al., 2017.

Chapter 6. Results 4: Structural and functional studies of the Ndc10/Scm3 interaction

The largest subunit of CBF3, Ndc10, has been shown to interact with the Cse4-specific histone chaperone Scm3, which led to the hypothesis that CBF3 is involved in recruiting and/or loading of the Cse4-nucleosome to the centromere. Structural data is only available for the NTD of Ndc10, as full-length protein is too unstable for crystallisation. Here interaction studies between Scm3 and Ndc10, as well as the full CBF3 complex are described. Additionally, efforts were made to gain structural insights into this interaction and the previously unknown CTD of Ndc10.

6.1 Expression and purification of Ndc10 constructs

Previous studies (Perriches, 2014) showed that Ndc10 expressed in bacteria readily degrades into the NTD (1-551) and CTD (551-end). Whereas the NTD is generally stable and has been crystallised (Perriches and Singleton, 2012); Figure 50), the CTD is less stable and if expressed and purified shows various degradations, with one at around ~25kDa being the most prevalent (Figure 51A). As digestion of the N-terminal sumo-tag did not lead to a different migration behaviour of this degradation product, it was assumed to be a N-terminal truncation. Edman sequencing, conducted by Alta Bioscience (Cambridge), confirmed this assumption by revealing the start residue to be L746. A construct from this residue to the very C-terminus has a size of ~25kDa, which corresponds to the mass observed on SDS-PAGE. Expression and purification of this construct, named CTDshort, yielded a stable and highly-expressed product (Figure 51B), which was deemed more suitable for structural studies than the unstable full CTD. It was also attempted to express the other half of the CTD, named centralCTD (residues 551-745). This, however, did not lead to an expression at any of the various conditions tested, which might not be surprising given that most of this region is predicted to be unfolded (Perriches, 2014).

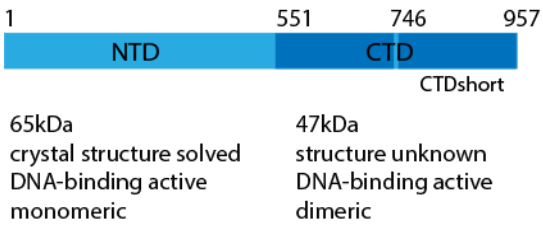

Figure 50: Ndc10 constructs
Schematic of the Ndc10 protein to give an overview of the different constructs used in this thesis.

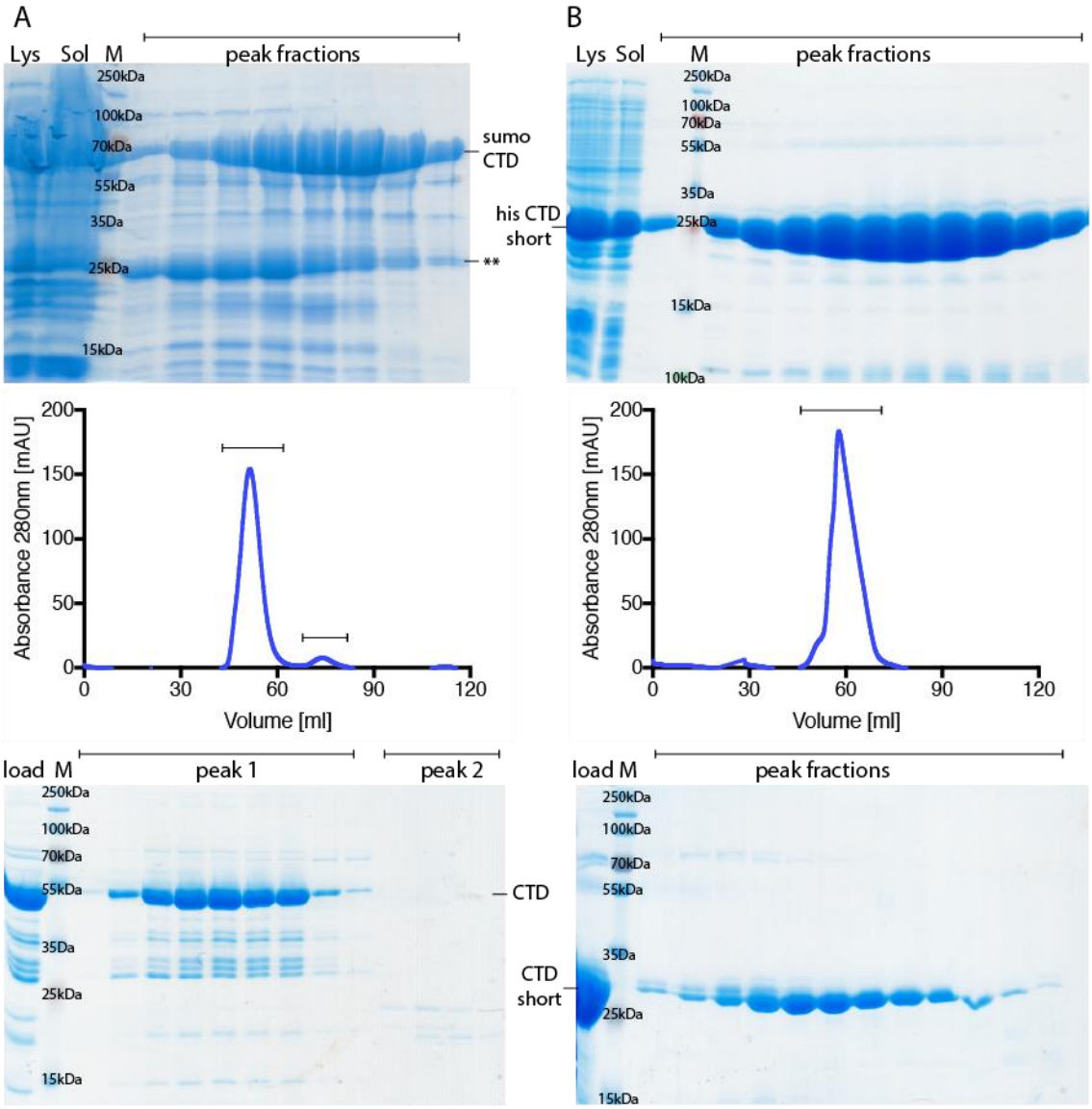

Figure 51: Purification of Ndc10 CTD and CTDshort

Upper panel: SDS-PAGE/Coomassie stain of lysate (Lys), soluble fraction (sol) and peak fractions from the initial affinity purification of CTD (**A**) and CTDshort (**B**). **... prominent ~25kDa degradation band, which was identified as the CTDshort construct, shown in **B**. Middle panel: typical size exclusion chromatography elution profile of CTD (**A**) and CTDshort (**B**). Lower panel: corresponding SDS-PAGE/Coomassie stain of the input (load) and peak fractions.

6.2 Interaction studies of Ndc10 and Scm3

Camahort et al., 2007 have been first to show an interaction between Scm3 and Ndc10 through CoIP, and Cho and Harrison, 2012 presented an *in vitro* pull-down

assay with the *K.lactis* homologue. To test if this *in vitro* interaction can be reproduced with budding yeast proteins, ITC studies were conducted with Ndc10 CTD and Scm3 full-length protein. Scm3 was expressed and purified from bacteria cells for this purpose (Figure 52A-C).

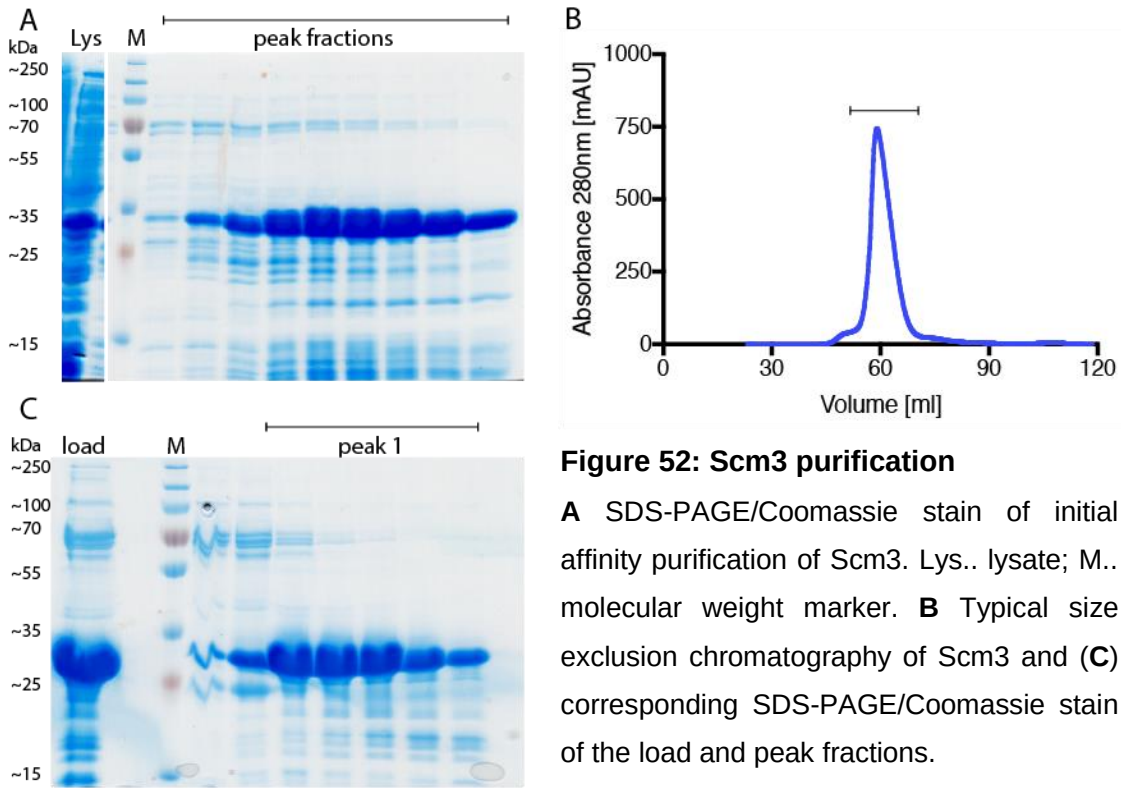

Figure 52: Scm3 purification

A SDS-PAGE/Coomassie stain of initial affinity purification of Scm3. Lys.. lysate; M.. molecular weight marker. **B** Typical size exclusion chromatography of Scm3 and (**C**) corresponding SDS-PAGE/Coomassie stain of the load and peak fractions.

As expected, ITC showed an interaction between Ndc10 CTD and Scm3 with a K_D of 187.9nM ± 37.7nM and a N of 0.6, suggesting a dimer of Ndc10 binding to a monomer of Scm3 (Figure 53A). Interestingly, if repeated with Ndc10 CTDshort, no interaction could be recorded, limiting the binding interface to residues 551-745 (Figure 53B). These results confirm well with previous studies, which showed that an equivalent *K.lactis* construct binds to Scm3 (Cho and Harrison, 2012). To get a better understanding of which part of Scm3 is important for this binding, a series of mutant peptides were designed and tested with ITC. Peptides were chosen from the extreme N-terminus (see Table 2), taking in account existing peptide array data and sequence conservation of Scm3 (Perriches, 2014). Whilst wild-type, mutant 1 and 2 peptides showed an interaction with Ndc10 CTD, mutant 3 and 4 were impaired in their ability to bind Ndc10 CTD. As Mutant 4 is a combination of all other mutants, residues L23 and L24 seem to be important for the interaction (Figure 53D-H).

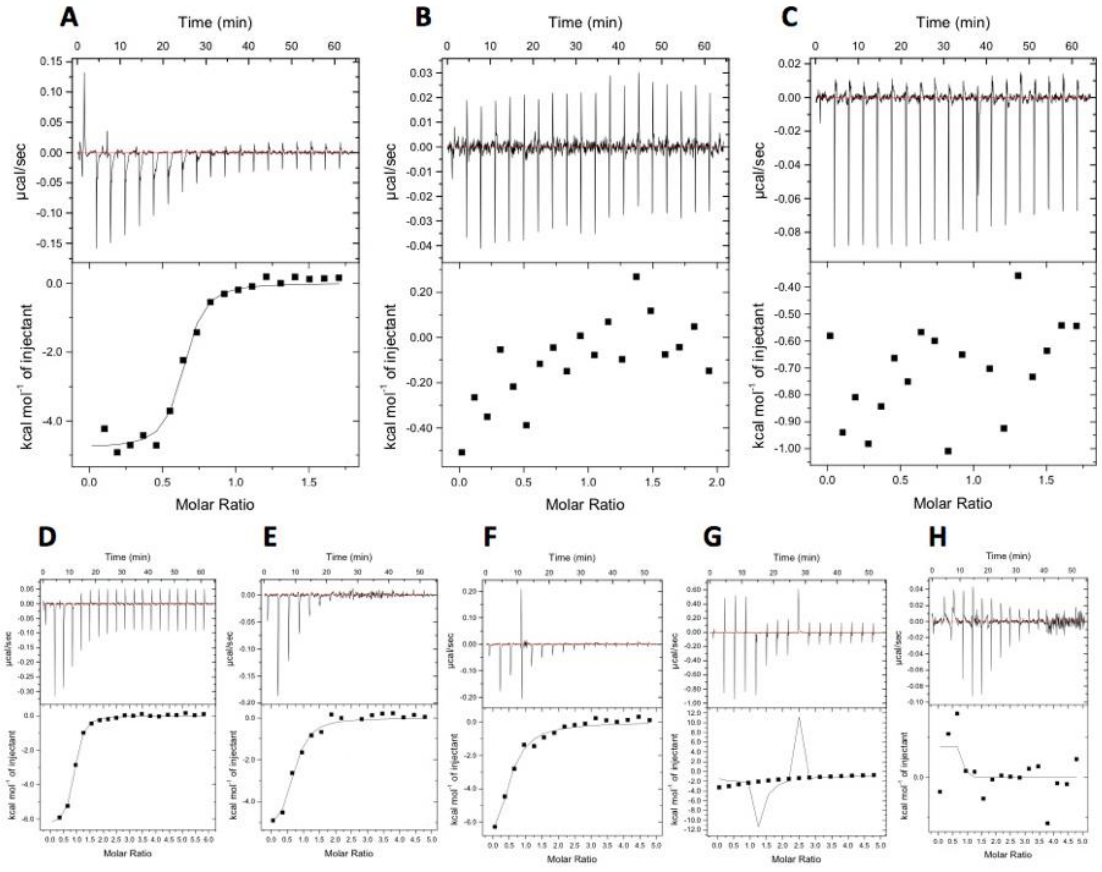

Figure 53: ITC studies of Ndc10/Scm3 interaction

Proteins were individually expressed and the buffer was matched by gel filtration with ITC buffer. Peptides were synthesised by the peptide unit and dissolved in ITC buffer. Measurements were done on a MicroCal iTC200 (Malvern), with 20 injections per run. Protein/peptide concentration in the syringe was 10 times higher than the protein concentration in the cell. Data was analysed using the Origin software. Ndc10 CTD and Scm3 (**A**) showed an interaction with a calculated KD of 187.9nM ± 37.7nM and N of 0.611, whereas Ndc10 CTDshort and Scm3 (**B**) did not interact. C shows the buffer control. **D-H**: ITC measurements with Ndc10 CTD and Scm3 peptides (see Table 2). Wild-type sequence peptide (**D**), mutant 1 (**E**) and mutant 2 (**F**) interact with Ndc10 CTD, but this interaction is abolished with mutant 3 (**G**) and mutant 4 (**H**).

In agreement to ITC measurements, analytical size exclusion chromatography with a mixture of Scm3 and Ndc10 CTD or CTDshort showed a co-elution only between Scm3 and Ndc10 CTD, but not with the CTDshort construct (Figure 54A and B).

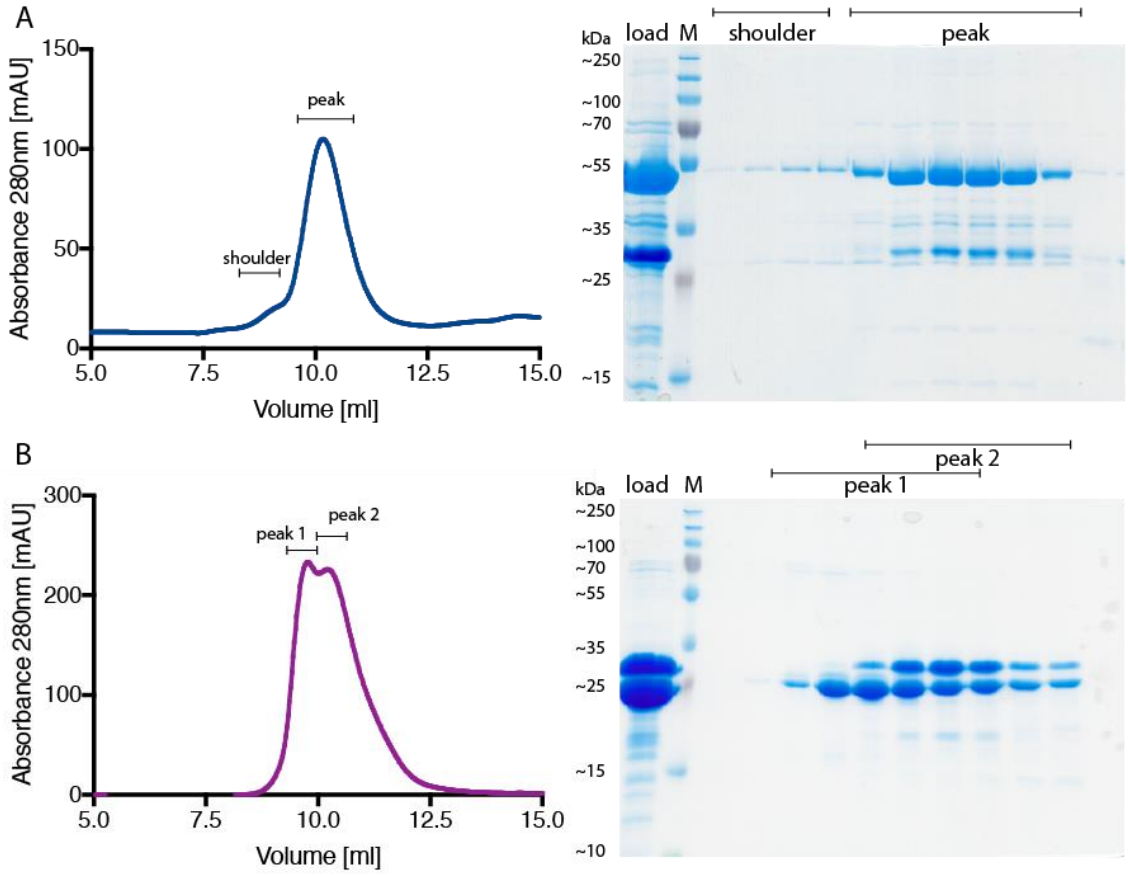

Figure 54: Analytical size exclusion chromatography with Scm3/Ndc10

Interaction between Scm3 and Ndc10 CTD (**A**) and Ndc10 CTDshort (**B**) tested with analytical size exclusion chromatography. Left panel: size exclusion profiles showing that Scm3 co-elutes with Ndc10 CTD on a S200 10/300 (GE Healthcare) column, but not with CTDshort on a S75 10/300 (GE Healthcare) column; right panel: corresponding SDS-PAGE/Coomassie stain. Proteins were mixed in an equal molar ratio and incubated for 10min on ice before loading.

6.3 Crystallisation trials

6.3.1 Ndc10 CTDshort

Given the instability of the CTD, crystallisation to elucidate the structure of the unknown C-terminal part of Ndc10 was only attempted with the CTDshort construct. It was concentrated straight after size exclusion chromatography and exhaustive crystal trials were conducted at different incubation temperatures, as well as with different protein concentrations (see 2.9.1 for details). However, no crystals were

obtained. Furthermore, efforts to determine the oligomeric state of this construct remained ambiguous and both SEC-MALS and analytical ultra-centrifugation (AUC) indicated multiple states and aggregations. Crystallisation trials were therefore ended.

6.3.2 Ndc10 CTD with NTD of Scm3

After confirming the interaction between the CTD of Ndc10 and Scm3, as well as demonstration that the very N-terminus of Scm3 is important, a truncated construct of Scm3 was designed. This was used to test the interaction and facilitate crystallisation trials with both interaction partners present. The binding to Scm3 might help to stabilise Ndc10 CTD and enable crystallisation.

Co-expression and purification of his-tagged Scm3 NTD and untagged CTD was successful and both proteins could be detected after initial affinity purification (Figure 55A). A substantial amount of CTD degradation into CTDshort was detectable, speaking somewhat against a strong stabilisation of CTD when bound to Scm3. However, size exclusion chromatography clearly showed a co-elution of the two interaction partners, with a ratio of two Ndc10 CTD to one Scm3 NTD, as judged by band intensities (Figure 55B and C).

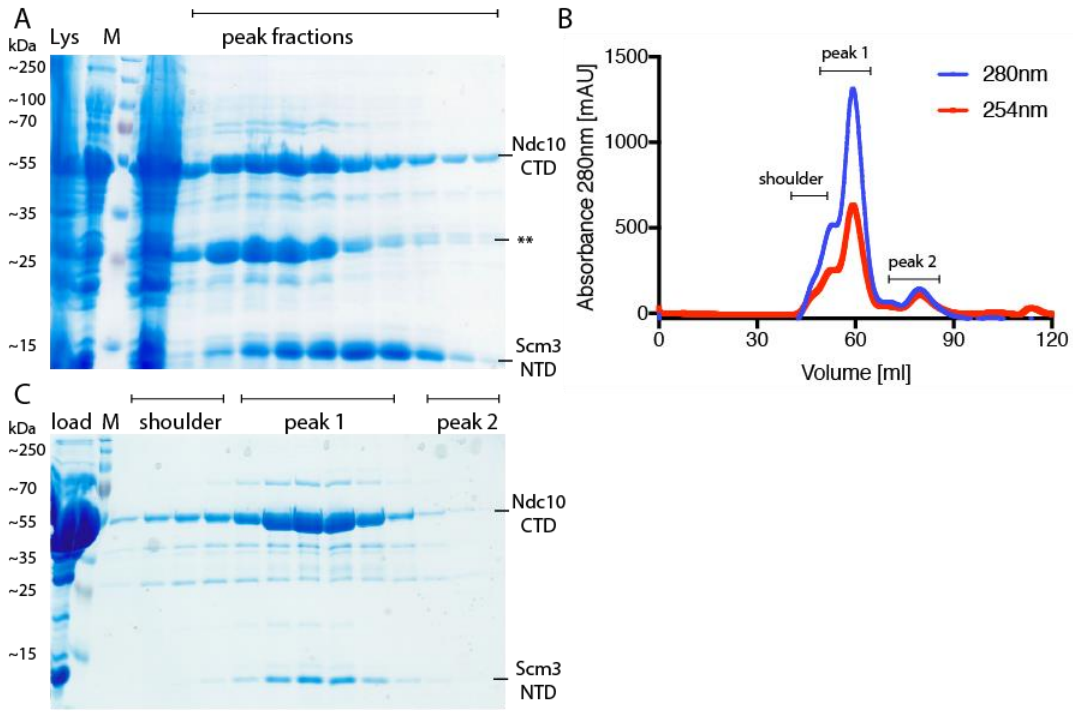

Figure 55: Co-purification of Ndc10 CTD and Scm3 NTD

A SDS-PAGE/Coomassie stain of initial affinity co-purification of Ndc10 CTD and Scm3 NTD. Lys... lysate; M... molecular weight marker; **… prominent degradation product identified as CTDshort. Typical size exclusion chromatography (**B**) and corresponding SDS-PAGE/Coomassie stain of the load and peak fractions (**C**).

Co-eluted proteins were concentrated and crystallisation trials were conducted (see 2.9.1 for details). As drops from the initial screens with 2.5 and 5mg/ml protein concentration, stayed mostly clear, concentration was raised to 6.8mg/ml. Further concentration was not possible due to aggregations becoming apparent. To further increase protein concentration, plates were set up with a 3:1 (protein to reservoir solution) ratio, as well as the more commonly used 1:1 ratio. Several initial hits were recorded in the 3:1 ratio drops on the PACT screen, with the best condition being 100mM Bis-Tris Propane (pH 6.5), 200mM sodium bromide and 20% PEG 3350 (Figure 56A). Other conditions included the same buffer and precipitant, but different salts, including 200mM sodium citrate, formate, acetate or K/Na tartrate (Figure 56B). All crystals were needles, and they appeared after ~24hrs at room temperature, but not at 4°C. Crystals only grew for about another 24hrs, before starting to bend.

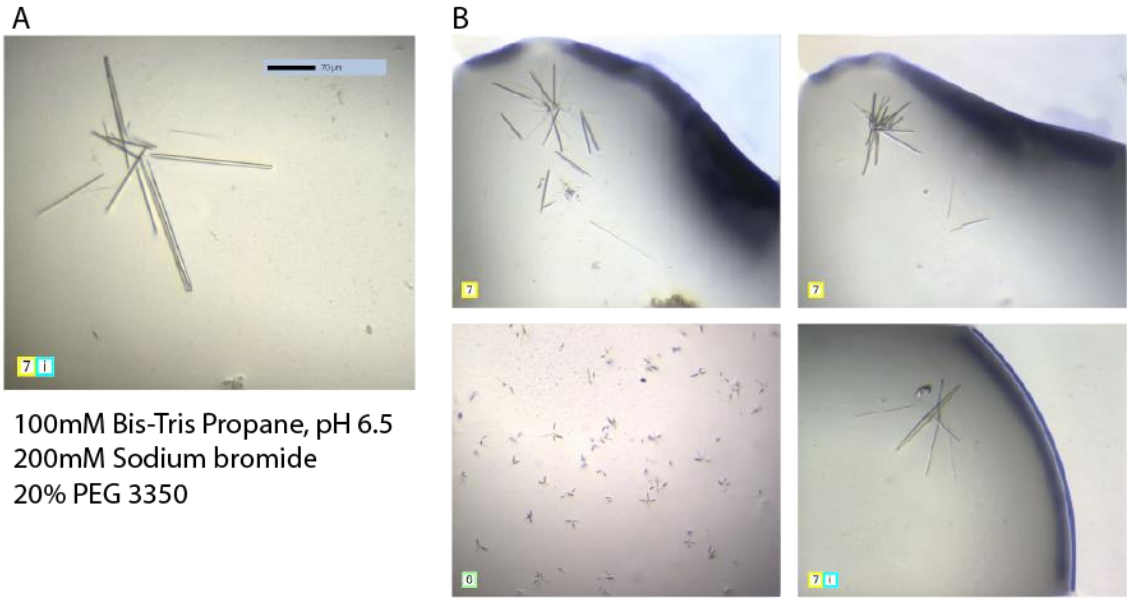

Figure 56: Ndc10 CTD/Scm3 NTD initial crystals

A Best initial hit, with specified condition. **B** Other hits, with similar conditions than **A** but different salt types including 200mM sodium citrate (top left), formate (top right), acetate (bottom left) and potassium/sodium tartrate (bottom right).

Optimisation of crystals was nonetheless attempted, by screening for optimal pH, salt type and concentration, as well as different PEGs and their concentrations. Crystals could be somewhat improved, however, diffraction was poor with spots being recorded only to about 15Å (Figure 57A). To test if crystals still contained both proteins, they were gently washed with reservoir solution and run on a SDS-PAGE, revealing that Ndc10 CTD was completely degraded and only Scm3 NTD and some slightly bigger degradation products were still present (Figure 57B). Crystal trials were therefore ended due to instability of Ndc10 CTD.

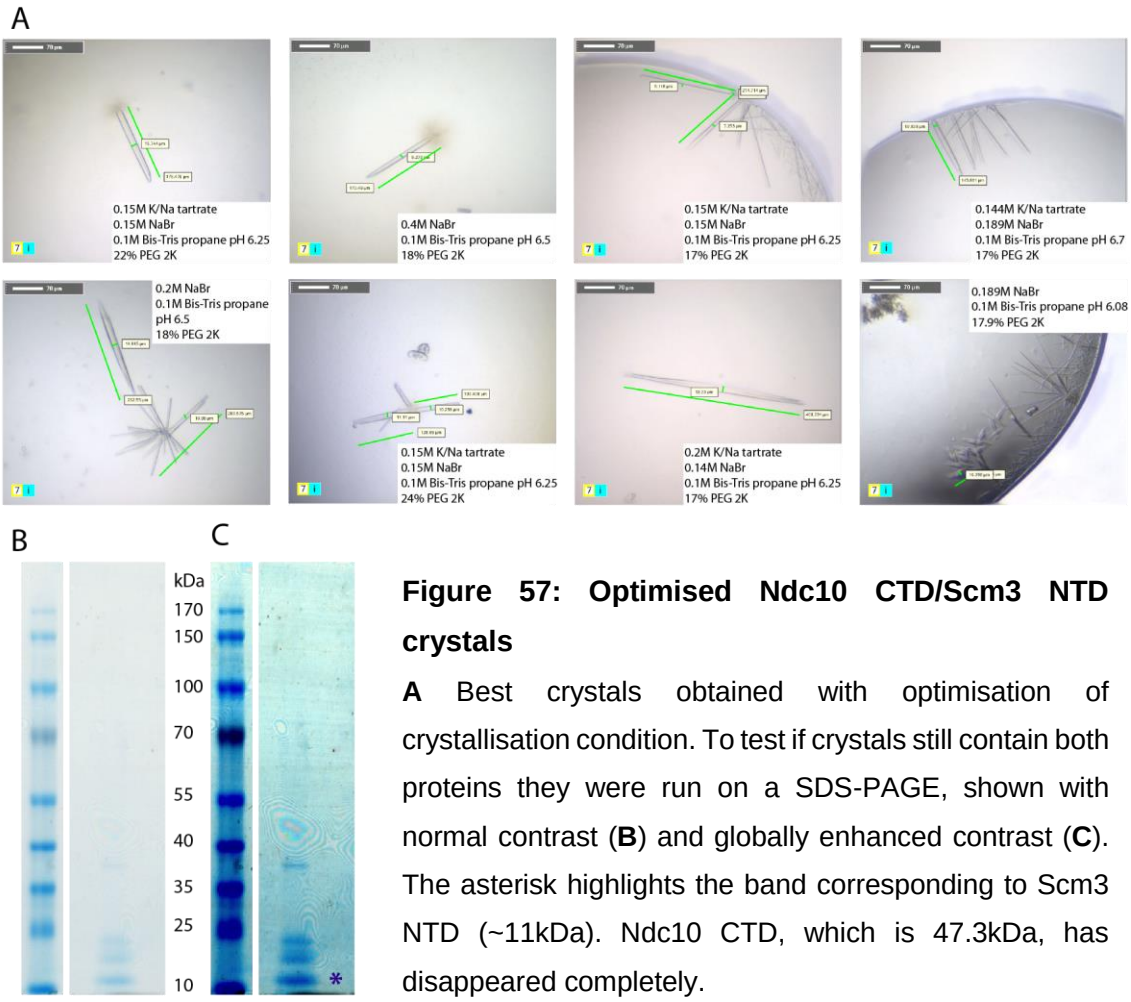

Figure 57: Optimised Ndc10 CTD/Scm3 NTD crystals

A Best crystals obtained with optimisation of crystallisation condition. To test if crystals still contain both proteins they were run on a SDS-PAGE, shown with normal contrast (**B**) and globally enhanced contrast (**C**). The asterisk highlights the band corresponding to Scm3 NTD (~11kDa). Ndc10 CTD, which is 47.3kDa, has disappeared completely.

6.4 Interaction studies of CBF3 and Scm3

After verifying the interaction between Ndc10 CTD and Scm3, it was attempted to test if this interaction can also be seen between Scm3 and Ndc10 full length in the context of the CBF3 complex. Purified full length Scm3 and tagged CBF3 complex were used for a pull-down assay with Streptavidin beads. No clear interaction was

seen between the two. The same amount of Scm3 was detected in samples with full CBF3, CBF3 core (no Ndc10 present) as well as the negative control with Scm3 only (Figure 58A). To see if a complex of Scm3, Cse4 and H4 histones, is needed for interaction, analytical size exclusion chromatography between such a complex and CBF3 was conducted. This also showed no interaction, as full CBF3 complex, core complex, Scm3 and the histones, seem to elute separately (Figure 58B). Interestingly Scm3 and the histones also dissociated. Further analyses, however, are necessary to fully interpret these preliminary results.

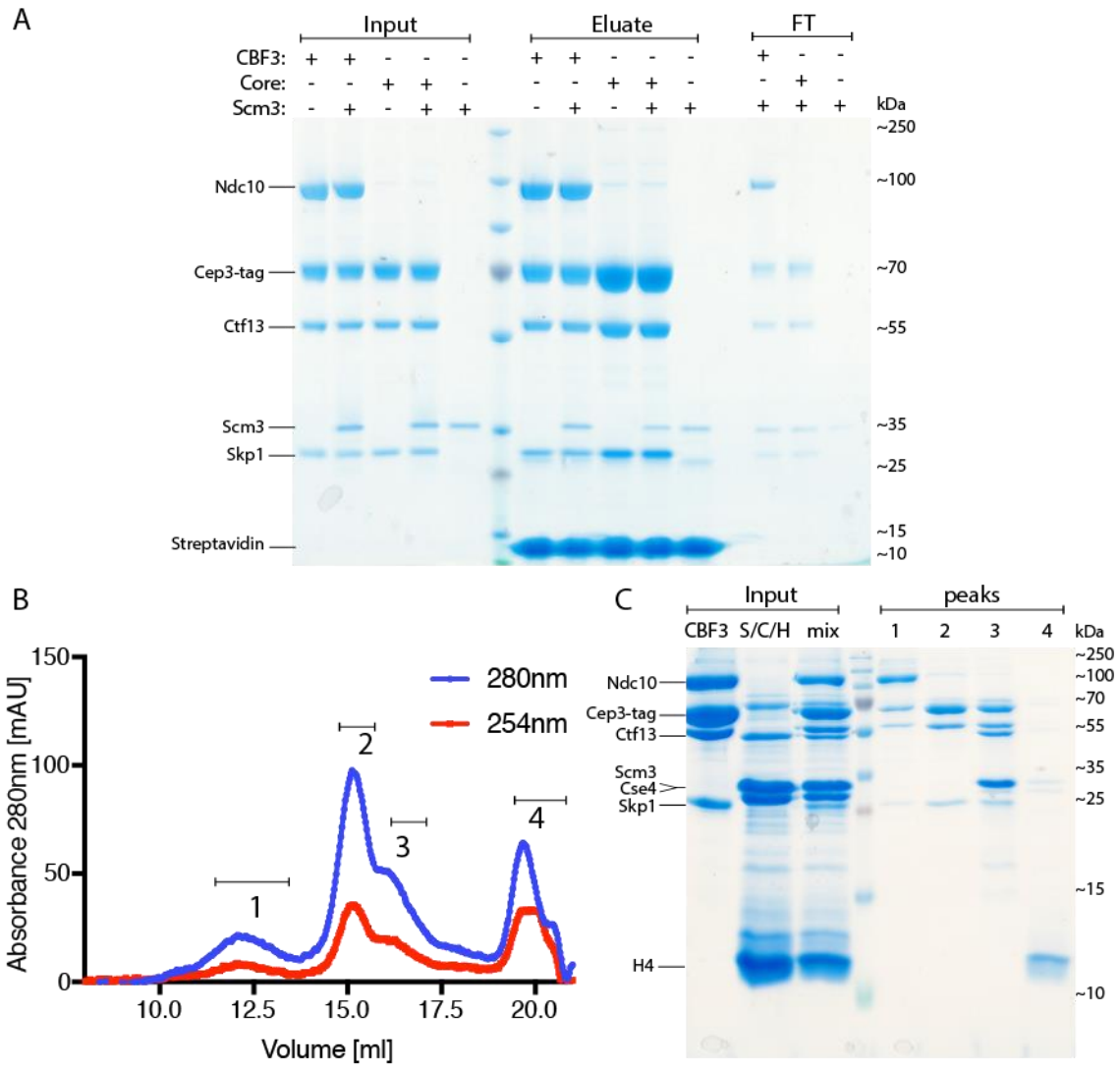

Figure 58: Interaction studies between Scm3 and CBF3 complex

A SDS-PAGE of a pull-down assay with purified Scm3 and CBF3, as well as with core and Scm3 alone as controls. Input is shown on the left side of the gel, eluate in the middle and flow-through (FT) on the right. **B** Size exclusion chromatography elution profile of CBF3 mixed with a Scm3/Cse4/H4 complex (S/C/H). Protein elutes in four major peaks, and no

interaction can be seen between CBF3 and Scm3, nor with Scm3/Cse4/H4 complex. The latter seems to dissociate during the chromatography. **C** Corresponding SDS-PAGE gel, showing the input samples on the left and the four peak fractions on the right. Due to low yield, peak fractions were pooled as indicated in (**B**) and concentrated to allow for Coomassie detection. Peak 1: CBF3 full complex; peak 2: core complex; peak 3: Scm3 and core (still present as peaks 2 and 3 are overlapping); peak 4: mostly H4 but also small amounts of Scm3 and Cse4.

Chapter 7. Discussion

7.1 CBF3 expression strategy and complex assembly

Expression and the role of Sgt1: The co-expression and purification strategy of the whole CBF3 complex described in this thesis should provide a useful tool for enabling further structural and functional studies not only of CBF3 on DNA, but also bigger budding yeast kinetochore assemblies. To my knowledge, it is the first strategy described with a high enough yield for structural studies. Successes have been made before with *K.lactis* CBF3 core complex, which could be expressed and purified from insect cells with co-expression of Sgt1 (Cho and Harrison, 2012). This and the finding presented here that simultaneous co-expression of Sgt1 considerably improves the yield, as well as the inability to express and co-purify CBF3 without Sgt1 co-expression in insect cells, emphasise previous studies showing the involvement of the Sgt1-Hsp90 chaperone pathway (Kitagawa et al., 1999, Bansal et al., 2004, Rodrigo-Brenni, 2004, Lingelbach and Kaplan, 2004). Interestingly, co-expression without Sgt1 yields an excess of free Cep3, indicating either that Sgt1 is necessary for the formation of Skp1/Ctf13 heterodimer or for the interaction between the heterodimer and Cep3. A recent crystal structure of a Skp1/Sgt1 complex has shown that both Cep3 and Sgt1 interact with Skp1 via overlapping interfaces (Willhoft et al., 2017). Thus, Sgt1 must dissociate before Cep3 can bind, which is further supported by the finding that expression of the Skp1/Ctf13 heterodimer does not co-purify with the simultaneously over-expressed Sgt1. Accordingly, purified Skp1/Ctf13 and Cep3 can form the core complex without Sgt1 present, as shown with an *in vitro* pull-down assay. Hence it is more likely that the Sgt1/Hsp90 chaperones are only involved in the formation of the heterodimer, which then in turn is able to bind to Cep3 and subsequently Ndc10 on its own. Despite the assumption that the heterodimer is very unstable and readily degraded if not incorporated into the more stable core or full complex, co-expression of the heterodimer yield is surprisingly high. This could stem from the recombinant expression strategy used and further studies are needed to address this. It would also be interesting to compare Skp1/Ctf13 expression without Sgt1 over-expression, which should yield significantly less protein.

Phosphorylations on CBF3: Despite earlier studies stating phosphorylation to be essential for complex formation and DNA-binding (Lechner, 1991, Kaplan, 1997) no direct evidence was found here. However, phosphatase inhibitors in the lysis buffer considerably impact the yield, indicating for some role of phosphorylation, which could be direct or indirect. Potential phosphorylation sites have been found on all four subunits (Holt et al., 2009, Kao et al., 2014, Kuilman et al., 2015) and the effect of phospho-mutants of some of these (Cep3 S575, Skp1 S162 and T177, as well as the Skp1 37-64 loop) were tested here. All showed no impact in complex formation, a finding consistent to a later study challenging the assumption of the need for phosphorylations (Stemmann et al., 2002). This study, however, only showed that complete dephosphorylation of the Skp1/Ctf13 heterodimer did not impact further interaction with Cep3 and Ndc10. It cannot be excluded that the Skp1/Ctf13 heterodimer formation itself is dependent on phosphorylation or that Ndc10 might need to be phosphorylated to be able to be bound by the other subunits. Cep3, which has been expressed in bacteria for this study and therefore cannot be phosphorylated can thus be excluded. Accordingly, phospho-mutants of the known phosphorylation site of Cep3 (S575; Holt et al., 2009, Strecker et al., 2016) had no influence on complex formation. An interesting potential site for Ndc10 phosphorylation is Y706, which has been found in a global screen for potential targets of Cdk1 and there is also evidence of a phosphorylation on this site from mass spectrometry results from this thesis. *In vitro* pull-down assays of the Ndc10 interaction with the core were carried out with Ndc10 which was co-purified with the other CBF3 subunits and subsequently separated during purification. It therefore could have been phosphorylated at this or other sites. The Y706 also lies within the CTD of Ndc10, which on its own did not show any interaction with the core complex in pull-down assays, whereas the NTD did. Both were expressed in bacteria and therefore were not phosphorylated. These are, however, mere indications and more thorough studies will be needed to clarify if direct phosphorylation of any of the subunits plays a part in complex assembly.

All findings presented here, support the suggested model of further complex assembly, which is an interaction between Cep3 and Skp1/Ctf13 forming the core complex and subsequent interaction with Ndc10. The core complex has been shown to be very stably bound, whilst Ndc10 can simply dissociate and re-associate.

Ndc10 troubles: Although Ndc10 could be co-purified and also formed a stable full complex with the core subunits, the particles visualised with negative stain EM were unsuitable for further structural studies and Ndc10 seemed to mostly dissociate from the core and generally be unstructured. This could be an intrinsic characteristic of the Ndc10 protein or a result of the purification or sample preparation process. Ndc10 has been described as a platform for many other inner kinetochore proteins to assemble, as it can interact with a number of different proteins (Cho and Harrison, 2012). Given this and the high predicted disorder, the unstructured density might not be very surprising. Furthermore, it is a very unstable protein on its own, as has been seen during the purification process. It also incorporates a degron motif at the very C-terminus, which upon exposure through unfolding of the protein leads to rapid degradation (Furth et al., 2011, Alfassy et al., 2013). However, removal of this degron motif by C-terminal truncations did not improve recombinant expression, nor the quality of particles. The NTD alone, on the other hand, seems to be too unstably bound to the core, indicating a longer construct is needed for proper interaction. This could also be verified with both crosslinking/mass spec, as well as a peptide array. Both showed that most of the interaction involves the NTD, but also some contribution of the CTD. One should also keep in mind that the NTD on its own is monomeric as the dimerization domain is excluded and this could impact its binding stability to the core complex. Future studies should concentrate on exactly identifying the stable product seen with limited proteolysis to test if this construct improves particles for structural studies.

The above-mentioned crosslinking/mass spectrometry results also agree well with previous studies concluding that Ctf13 forms the central part of the complex (Russell, 1999), as three of the four crosslinks seen on Ndc10 are with Ctf13. None are seen with Skp1, which would indicate Ndc10 binds distal of Skp1. If this is the case the Skp1 37-64 loop would remain solvent exposed and could therefore still be phosphorylated or dephosphorylated to ‘regulate’ DNA binding. Indeed, initial EMSA studies comparing the DNA binding capability of the full CBF3 complex before and after lambda phosphatase treatment do show an increased binding once dephosphorylated. The fourth crosslink of Ndc10 is to the Gal4-domain of Cep3. As Cep3 is a dimer, this could either be to the already stabilised Gal4-domain bound by Skp1 and Ctf13, or to the remaining free, DNA-binding active Gal4-domain. If the latter is involved, one could assume a stabilisation of Ndc10 interaction with the core

upon DNA binding, as this would stabilise this Gal4-domain, strengthening the overall interaction of all four subunits. Indeed, if one compares a co-purification of wild-type and coreSΔ the latter pulls down more Ndc10 consistently. The reason for this is believed to be that it is DNA-binding active, rather than any direct influence of the Skp1 37-64 loop, as after this initial difference both wild-type and coreSΔ behave identically. Similarly interesting is that this crosslink links to the very C-terminus of Ndc10, the same stretch shown to interact with the core by peptide array and also the same stretch harbouring most of the degron motif. It is possible that the latter becomes inaccessible for the proteasome machinery upon being bound by the core subunits, leading to an increased stability of Ndc10 within the complex. Indeed, it was only possible to co-purify Ndc10 with the other subunits, whilst expression and purification on its own in budding yeast did not yield any protein. Whilst being an interesting assumption, further biochemical studies or structural data of the full CBF3 complex are needed to clarify these points.

Another reason for Ndc10 being unstructured could be that an interaction partner is needed, with the centromeric DNA being the most likely candidate. Although both the core and full complex were DNA-binding efficient, the binding proved to be too weak to successfully prepare negative stain EM grids and therefore achieve any structural data. One likely reason is the relatively high salt concentration needed to prevent the complex from aggregation, but possibly hindering DNA interaction. Further buffer optimisation, maybe including surfactants and screening of different salts, might be the key to successfully reconstitute the DNA/CBF3 complex for structural studies.

The first study to describe an *in vitro* reconstitution of endogenous CBF3 with centromeric DNA also found the need for an assembly factor, as addition of either a small amount of yeast or *E.coli* lysate or purified casein was needed to successfully reconstitute the complex. Here, the addition of casein did not improve DNA binding and most likely the effect of these assembly factors seen in this study simple came from the prevention of aggregation of CBF3 at the low salt concentrations used in the assays, as also suggested by the authors themselves (Lechner, 1991).

Revised model of CBF3 assembly: CBF3 assembly starts with the formation of the Skp1/Ctf13 heterodimer, which needs the action of the Sgt1/Hsp90 chaperones

interacting with both Skp1 and Ctf13. The chaperones dissociate after formation of the heterodimer, which becomes stable whereas Ctf13 alone is not (Figure 59A).

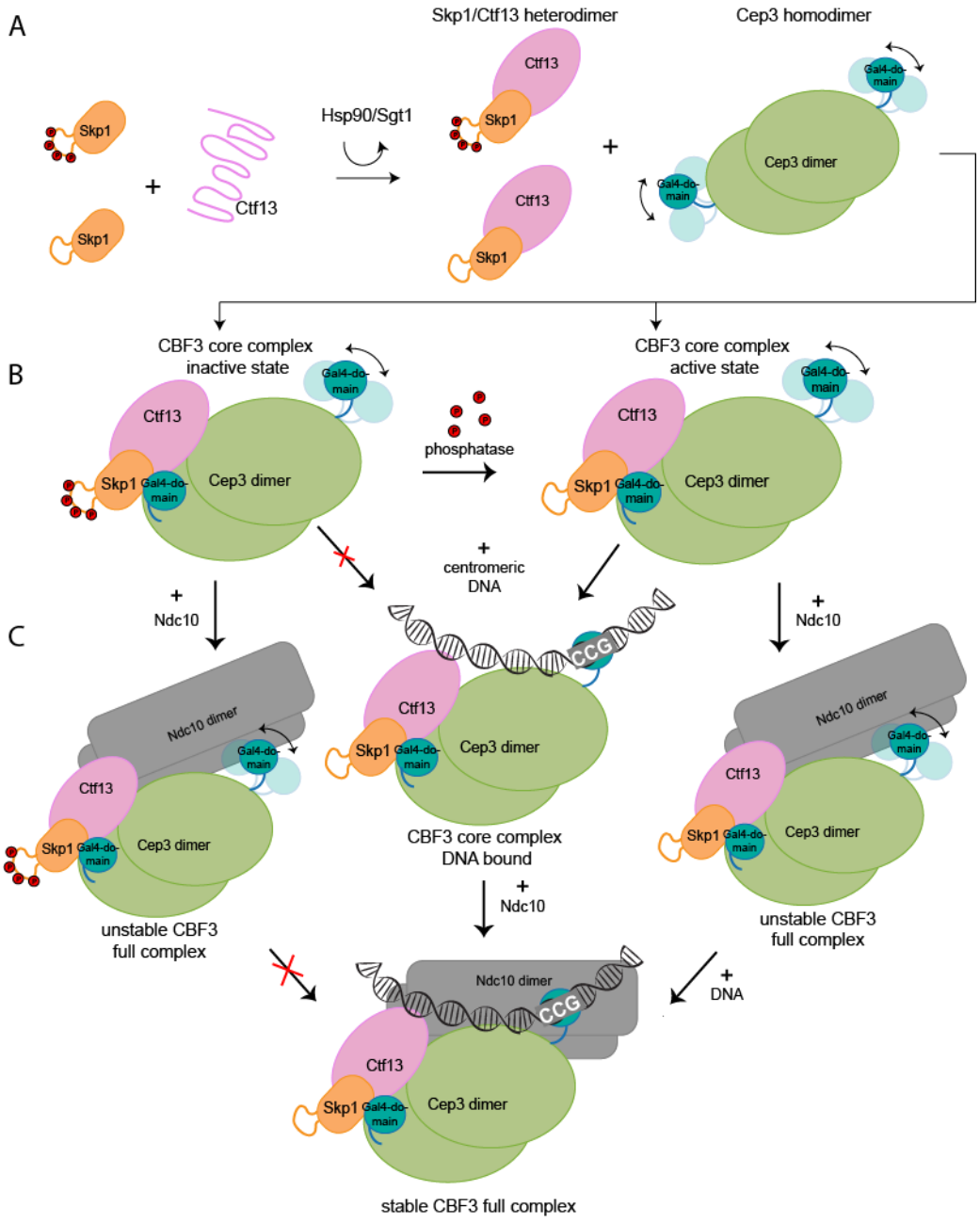

Figure 59: Revised model of CBF3 complex assembly

Assembly starts with the formation of the Skp1/Ctf13 heterodimer in a Sgt1/Hsp90 dependent manner (**A**). The Skp1/Ctf13 heterodimer can then bind to a Cep3 homodimer (**B**) and form the core complex, which can interact with either Ndc10 or DNA and ultimately lead to the formation of the full CBF3 complex bound to centromeric DNA (**C**). DNA binding of the core as well as the full CBF3 complex is dependent of the phosphorylation status of the 37-64 Skp1 loop. *Note: Part of this figure was adjusted from Leber et al. 2018.*

The next step is the spontaneous binding of the heterodimer to Cep3 forming the CBF3 core complex (Figure 59B). This happens independently of the phosphorylation status of the Skp1 37-64 loop or any assembly factors and leads to the stabilisation and therefore inactivation of one Gal4-domains. *In vitro* this core complex is able to either bind to Ndc10 forming the full CBF3 complex or to centromeric DNA, whereas only the latter is dependent on dephosphorylation of the Skp1 37-64 loop. Therefore, the order of the last two events, the binding of Ndc10 and to centromeric DNA, remains unclear and might even be interchangeable (Figure 59C). The full complex without DNA, however, forms a rather loose association where Ndc10 can readily dissociate, which probably stems from the high disorder of Ndc10 without being bound to DNA or other kinetochore proteins and might also be influenced by the flexibility of the remaining free Gal4-domain.

7.2 CBF3 core cryoEM structure

The CBF3 core complex structure presented in this thesis is a great example of the power of cryoEM. Due to low yields and its tendency to aggregate, crystallisation of CBF3 core would have been very difficult and time consuming at best, and impossible at worst. Crystallisation needs the orderly arrangement of the protein/protein complex to form a crystal lattice. Proteins which tend to aggregate severely hinder this, as they randomly attach to each other. A limitation for the achieved resolution of the cryoEM reconstruction was most likely the thick ice needed to avoid degradation of the protein when interacting with the air-water interface. This is a common problem with many protein samples and it may be possible to improve on this in future studies by the use of surfactants or addition of interaction partners such as Ndc10 or the centromeric DNA.
The tight overall interface of the core structure, explaining its stability, is the first structural insight into the interaction between the subunits, of the shape of Ctf13 and has elucidated some peculiar features, like the stabilisation of one Gal4-domain and a new mode of F-box binding, all of which is discussed in detail below.

Ctf13 structural insights: Although the resolution obtained did not allow for building a full atomic model of the previously unknown Ctf13 subunit, it unambiguously identified as an F-box protein containing LRRs (FBXL) like Skp2. This also fits with the role of Sgt1 in assembly of the heterodimer as studies have shown that it

specifically targets proteins containing LRRs (Taipale et al., 2014, Stuttmann et al., 2008). Ctf13 has, however, diverged from typical FBXL proteins, incorporating an extended helical domain after the N-terminal F-box as well as a helical insertion in the LRRs, which functions as a cap on the F-box distal end of the structure. This is atypical for proteins with LRRs, as they usually form a cap with their very C-terminal domain. Accordingly, the sequence of Ctf13 has diverged significantly both in the F-box and the LRRs, in the case of the latter becoming almost unrecognisable. Intriguingly this has also been seen for Ndc10 NTD, which has an unexpected tyrosine recombinase fold unable to be predicted from the sequence alone. The LRRs of Ctf13 also represent another reused protein motif from unrelated cellular processes, such as the above-mentioned tyrosine recombinase fold of Ndc10 (DNA topology and transposition), the Gal4 DNA-binding domain from Cep3 (transcription), as well as the F-box of Ctf13 and Skp1 (ubiquitylation). Intriguingly all of these processes are ancient and widespread, speaking for CBF3 to be a relatively recent invention. This would fit with the hypothesis that point centromeres are a relatively young simplification of the more complex regional centromeres, emerged due to the lack of centromere competition in budding yeasts and actually representing a modification of the 2-micron plasmid partitioning system (Malik and Henikoff, 2002, Malik and Henikoff, 2009).

A new Skp1/F-box interaction: An interesting feature of the core structure is the way Skp1 binds the F-box of Ctf13, as it is different to all other known Skp1/F-box interactions. Intriguingly all of these structures were derived from proteins expressed in bacteria, whilst the phosphorylation status of Skp1 has been shown to affect the interaction with Met30, a F-box protein containing WD40 repeats (S162; Beltrao et al., 2012). Mutations in the two potential phosphorylation sites of Skp1 (S162 and T177) within the F-box binding part did not influence formation of the Ctf13/Skp1 heterodimer, as all phospho-mutants of Skp1 successfully co-purified Ctf13. The possibility that phosphorylation of these sites influences further complex formation like Cep3 binding, however, cannot be excluded and should be tested in the future.
Besides the two last helices of Skp1, which in all available structures facilitate F-box binding, an extended loop (residues 105-112) also contacts the F-box, as well as the LRRs of Ctf13 in the cryoEM structure. Two aspects of this are interesting, first this loop has been found disordered and solvent exposed in all known Skp1/F-box protein

structures solved so far, and second although Ctf13 only exists in budding yeasts this loop of Skp1 is conserved to higher eukaryotes. This could mean that there may be other FBXL proteins which bind in the same manner as Ctf13. Therefore, one could separate between two classes of FBXL proteins binding either with or without involvement of this loop. Another scenario would be that Skp1 can bind to one FBXL protein in two different conformational states, possibly involving some posttranslational modifications. The latter could represent a switch between inactive and active state, as has been described for the Skp1/Ctf13 heterodimer (Russell, 1999, Kaplan, 1997). Although an interesting possibility, not only for CBF3 but also for SCF ligases and the impact it could have on the positioning of the substrate within the SCF ligase, further studies are needed to test this hypothesis. For example, one could test if the Skp1/Ctf13 phosphor-mutants influence Cep3 binding with *in vitro* pull-down assays, or conduct *in vivo* interaction studies, like protein complementation assays with the different mutants.

Gal4-domain stabilisation: The Skp1/Ctf13 heterodimer binds to one Cep3 monomer in a way that includes contacts with the Gal4-domain. This leads to a stabilisation and inactivation of this Gal4-domain, whilst the other remains free. Therefore, CBF3 although containing a homodimer of Cep3 can only recognise one CCG triplet. A peculiar DNA-binding characteristic, which explains that only one CCG triplet is conserved in the point centromere. It also highlights the significance of structural biology to elucidate how proteins and protein complexes function, as one could have never predicted this behaviour from sequence alone.

The remaining flexible Gal4-domain could be visualised, albeit with low resolution, with sophisticated focussed 3D-classification algorithm, thus stressing the importance of good software in extracting all information of a cryoEM dataset. It does not seem to exhibit, however, distinct conformations, rather a transient movement, which is the reason for the low local resolution obtained. Therefore, the likely path of the DNA could not be predicted, and here again only structural data of the complex bound to DNA will elucidate the exact mode of DNA binding. Biochemical evidence of DNA binding is discussed in the next chapter.

7.3 DNA-binding of CBF3 core and full complex

Besides above-mentioned stabilisation and inactivation of one Gal4-domain by the binding of the Skp1/Ctf13 heterodimer, EMSA studies clearly showed that the core can bind centromeric DNA sequence-specifically and that the remaining free Gal4-domain is most likely responsible for this binding. This represents the first clear evidence of DNA-binding capability of the core complex without Ndc10 and stands against the proposal that all four subunits are needed to activate CBF3 (Sorger et al., 1995, Stemmann and Lechner, 1996, Kaplan, 1997, Espelin, 1997, Pietrasanta, 1999).

Intriguingly, the core can only bind after dephosphorylation of the 37-64 Skp1 loop, suggesting a form of regulation of DNA-binding previously unknown. Although Skp1 is conserved to higher eukaryotes, this loop is budding yeast-specific, as is CBF3 itself. *In vivo* studies showed that one can delete this loop (Skp1Δ) and cells are still viable (Connelly and Hieter, 1996) and mutation of the phospho-sites in this loop also results in viable cells (Stemmann et al., 2002). Accordingly, in this thesis it is shown that Skp1Δ can mimic the DNA-binding active state and does not influence core complex formation nor its structure. If the function of phosphorylation of this loop is only to inhibit DNA-binding then these results could be explained, as cells without the loop or phospho-sites would still be able to form the kinetochore and undergo mitosis. DNA-binding inhibition might, for example, only ensure that CBF3 binds to the centromere in a cell cycle regulated manner or hinders binding of CBF3 outside of the centromeric sequence. More stringent *in vivo* experiments are needed, however, to test these hypotheses and future studies should include the optimisation of the cell cycle regulation experiment of Skp1 phosphorylation in 5.1.3. Although the overall phosphorylation of Skp1 did not seem to change during the cell cycle, CBF3-specific changes could still take place and might not have been seen due to the unsuccessful Co-immunoprecipitation. One possible reason for this may have been the method used for cell disruption, as it has been found to also affect the complex integrity of recombinant expressed CBF3.

As structural data of the CBF3/DNA complex is still lacking, the details of DNA-binding and how the phosphorylation status of the Skp1 loop effects this interaction is not clear. Generally, one could suggest two possible ways: First, DNA could be bound by the free Gal4-domain and then lie across the core structure bringing it close

to the Skp1 37-64 loop and the charge of the four phosphorylations inhibits DNA binding. Alternatively, the dephosphorylation of the Skp1 loop could lead to a conformational change which allows stable DNA-binding. Intriguingly, the central channel of the overall core structure would accommodate a double strand DNA nicely with only a small movement of the Skp1/Ctf13 heterodimer to open the channel. The DNA-binding active coreSΔ, which mimics the dephosphorylated core in EMSA studies, however, does not seem to exhibit any movements and the central channel is of identical dimensions. Also, one would expect a rather strong interaction between the core and DNA if it would indeed lie tightly packed in the central channel and this does not seem to be the case in both the EMSA studies and during attempts to reconstitute the complex with DNA. Therefore, it seems more likely that DNA lies across pointing towards the Skp1 37-64 loop. It is difficult to predict the path of DNA, however, with the limited structural data, and future studies should concentrate on further optimisation of purification and/or reconstitution protocols, which should allow structural studies of the CBF3/DNA complex to elucidate the interaction and the possible effect of DNA bending and looping.

DNA binding in context of the nucleosome: Preliminary EMSA studies presented here showed that both core and full CBF3 seem to prefer to bind to free DNA rather than DNA wrapped around the nucleosome. Although further optimisation of these assays will be needed to clarify the exact effect of Cse4 and centromeric DNA, this result could reflect the fact that CBF3 is supposed to be the initial factor to recognise centromeric DNA and then leads to the recruitment of the centromeric nucleosome through other proteins, like Scm3. It could also mean that CBF3 needs naked DNA to be able to bind to the centromere and initiate kinetochore formation or that CBF3 DNA-binding is incompatible with a conventional left-handed nucleosome. Interestingly, a recent study has shown that the handedness of DNA wrapping depends on the supercoiling state of the DNA, and one can assemble both left and right-handed nucleosomes (Vlijm et al., 2017). Future studies could take advantage of this study and test if the binding is influenced by the handedness of the nucleosomes, as well as more stringently test if there is a preference to bind to naked DNA instead of DNA wrapped around nucleosomes.

The question how CBF3, the Cse4-nucleosome, Mif2, Cbf1 and possibly other kinetochore proteins can interact simultaneously with the short centromere remains.

An interaction between the core and DNA with the flexible Gal4-domain, however, would allow the core to bind with considerable freedom of movement relatively to the DNA and could position it just outside the nucleosome. The majority of the Cse4-nucleosome interaction could be facilitated through the subunit Ndc10, as has been previously suggested (Cho and Harrison, 2012). This would allow for a rather flexible arrangement, which could facilitate the Cse4-nucleosome assembly steps as well as could lead to unusual DNA topologies found at the centromere (Pietrasanta, 1999, Lawrimore et al., 2015, Diaz-Ingelmo et al., 2015).

7.4 Interaction between Ndc10 and Scm3

CBF3 is believed to recruit Cse4 specifically to the centromere via an interaction between the conserved histone chaperone Scm3 and Ndc10. Scm3 forms a trimeric complex with Cse4 and H4 in solution (Cho and Harrison, 2011, Zhou et al., 2011). This interaction does not include the N-terminus of Scm3, which is believed to be involved in the interaction with Ndc10. Here this finding is confirmed with interaction studies between the N-terminus of Scm3 and the CTD of Ndc10. The latter is the least characterised of Ndc10 and is involved in dimerization, DNA-binding and interaction to other kinetochore proteins like Scm3 and Bir1 (Cho and Harrison, 2012, Perriches, 2014). ITC studies could identify important Scm3 residues for the interaction with the Ndc10, however, it was not possible to stabilise a complex between the two for crystallisation studies. Hence, further optimisation will be necessary to find a shorter Ndc10 construct which is still Scm3 binding efficient but more stable for structural studies. Furthermore, mutational studies with the Scm3 protein rather than simply peptides should be conducted to verify the impact of these Scm3 residues on the interaction.

Surprisingly no interaction could be found between the full CBF3 complex and Scm3, although these results should be confirmed by additional experiments and replications. However, it would be possible that the unstructured nature of Ndc10 seen with negative stain EM is somehow hindering Scm3 interaction. Possibly, CBF3 needs to be bound to the centromere, before Ndc10 is able to recognise Scm3 and therefore helps to incorporate the Cse4-nucleosome by destabilising the Scm3/Cse4/H4 interaction. Both biochemical assays and structural studies should

be conducted to test this hypothesis and elucidate the role Ndc10 is playing in recruiting and possibly also incorporating the Cse4-nucleosome.

7.5 Recent publications

Just before submission of this thesis an article was published presenting a modified core structure with and without an equivalent to the NTD of Ndc10 by Zhang et al., 2018. The following section summarises their findings in respect to this thesis.

The authors describe a cryoEM structure of a construct lacking the Gal4-domains of Cep3 (identical to the core Gal4Δ construct used here) at a similar overall resolution. However, the Ctf13 density is better defined enabling an atomic model of Ctf13 to be built. They confirm findings presented here of the unusual F-box/Skp1 interaction, as well as the insertion in the LRRs capping the structure on the F-box distal side. However, they identified eight, rather than the seven LRRs reported here. It is interesting that the Ctf13 subunit is better defined in this construct rather than the wild-type, especially as the part of Ctf13 with lower resolution in our structure is directly in contact with the stabilised Gal4-domain, which is completely lacking in their construct. The other structural details are identical to the core structure presented in this thesis.

Similarly, they also co-expressed the core subunits in budding yeast itself with simultaneous expression of the assembly factor Sgt1. However, they report a difficulty in purifying the wild-type core complex with Cep3 containing the Gal4-domains. It is possible that this difficulty may have arisen from the lack of phosphatase inhibitors, as in their method section these are not mentioned. As described in this thesis, however, they are crucial for complex integrity.

In their case, DNA-binding can be seen without prior dephosphorylation, which again could stem from the lack of phosphatase inhibitor leading to dephosphorylation of the Skp1 loop during cell breakage. It would be interesting to test this by determining the phosphorylation status of their construct; however, the authors have not tried to assess this possibility but only state the difference. It should also be noted that whilst this thesis clearly shows that DNA-binding is facilitated by the Gal4-domain(s), their construct is lacking exactly these. Mutational studies have shown the importance of the conserved CCG triplet in the centromere, and it is clear that the Gal4-domain(s)

are responsible for this sequence-specific interaction (Ng and Carbon, 1987, Hegemann and Fleig, 1993, Bellizzi, 2007, Purvis and Singleton, 2008). However, other contacts are made with Cep3 and Ctf13 as has been shown by crosslinking studies (Espelin, 1997). The binding reported by Zhang et al., 2018 therefore must represent only these additional contacts. Furthermore, they map the DNA-binding to the central channel with mutational studies of Ctf13 and Cep3 residues, suggesting a model by which both Gal4-domains bind to DNA to the CCG and to the potential second TGT triplet. They argue that the core structure reported in this thesis is locked in an inactive state, and that the stabilised Gal4-domain needs to be freed to activate the complex. Indeed, the wild-type core structure presented here corresponds to the inactive DNA-binding state, however, unknown to the authors we excluded the possibility that the stabilised Gal4-domain must be set free for DNA-binding with the second cryoEM structure of the coreSΔ. This construct although DNA-binding active, still displays the same stabilisation and therefore inactivation of the Gal4-domain in the vicinity to the Skp1/Ctf13 heterodimer. However, it cannot be excluded that dephosphorylation, rather than deletion of the 37-64 Skp1 loop is unlocking this Gal4-domain. Equally it is possible that some part of the datasets both for wild-type and coreSΔ were comprised of particles with an unlocked or rotated Gal4-domain, which therefore would allow DNA contacts. If those particles do not average together, for example because this Gal4-domain is then more flexible, they could have been discarded in the data processing workflow. However, at no point in any classification step such a subset of particles was seen and this would also not explain the peculiar fact that there is only one conserved CCG triplet at the centromere. Furthermore, mutation of the TGT triplet has no effect on DNA-binding in EMSA studies presented here, which would not be expected if the stabilised Gal4-domain is specifically binding these residues, as suggested by Zhang et al. 2018. It should be noted that the idea of the second TGT triplet being bound by one Gal4-domain has only ever being suggested as a possibility and no data directly supports this binding. The crosslinking study cited by Zhang et al. 2018 only shows an interaction between these residues (TG) and the Cep3 protein as a whole, not the Gal4-domain especially (Espelin, 1997). Therefore, it was surprising that the authors stated that hypothesis as fact in their introduction, when it would be equally possible that any other part of Cep3 is in contact with these two bases. However, the mutational studies shown by the authors is speaking clearly for the DNA to be located in the central channel. A

possibility which in this thesis is also not excluded, even if the stabilised Gal4-domain is not involved in DNA-binding itself.

Excitingly, the authors could stabilise a complex of the core Gal4Δ and an equivalent of the NTD of Ndc10 for structural studies, an important achievement. Although lower in resolution they could dock the crystal structure of the NTD into the EM density and therefore describe the core binding site, as well as predict the DNA binding path. The interaction domain lies in the insertions of the Ctf13 F-box and the very N-terminus of Ndc10, supporting findings described here that the NTD is essential for the interaction. Interestingly, this brings the stabilised Gal4-domain in the wild-type structure presented here into close proximity of this binding domain. It will be interesting to see if the NTD interaction could be affected by the Gal4-domain by superposition of the two core structures upon release of the coordinates of the core Gal4Δ with Ndc10 NTD from the PDB/EMDB databanks. If this binding is not compatible, one would assume that indeed a conformational change needs to take place for further complex formation, possibly freeing the stabilised Gal4-domain. This could also explain the difficulties described here in stabilising Ndc10 on the core structure.

Even more recently another CBF3 structure with centromeric DNA has been deposited, although not yet released nor published. It will be exciting to see how this new structure binds DNA, and possibly how the differences between the wild-type and core Gal4Δ might be explained. All these efforts together should provide a more accurate picture of the function of CBF3 on the point centromere and are an important step forward in kinetochore research.

Reference List

ADAMS, P. D., GOPAL, K., GROSSE-KUNSTLEVE, R. W., HUNG, L. W., IOERGER, T. R., MCCOY, A. J., MORIARTY, N. W., PAI, R. K., READ, R. J., ROMO, T. D., SACCHETTINI, J. C., SAUTER, N. K., STORONI, L. C. & TERWILLIGER, T. C. 2004. Recent developments in the PHENIX software for automated crystallographic structure determination. *J Synchrotron Radiat,* 11**,** 53-5.

AEBI, U. & POLLARD, T. D. 1987. A glow discharge unit to render electron microscope grids and other surfaces hydrophilic. *J Electron Microsc Tech,* 7**,** 29-33.

AKIYOSHI, B., NELSON, C. R. & BIGGINS, S. 2013. The aurora B kinase promotes inner and outer kinetochore interactions in budding yeast. *Genetics,* 194**,** 785-9.

ALFASSY, O. S., COHEN, I., REISS, Y., TIROSH, B. & RAVID, T. 2013. Placing a disrupted degradation motif at the C terminus of proteasome substrates attenuates degradation without impairing ubiquitylation. *J Biol Chem,* 288**,** 12645-53.

ALLSHIRE, R. C. & KARPEN, G. H. 2008. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet,* 9**,** 923-37.

ALUSHIN, G. M., MUSINIPALLY, V., MATSON, D., TOOLEY, J., STUKENBERG, P. T. & NOGALES, E. 2012. Multimodal microtubule binding by the Ndc80 kinetochore complex. *Nat Struct Mol Biol,* 19**,** 1161-7.

ALUSHIN, G. M., RAMEY, V. H., PASQUALATO, S., BALL, D. A., GRIGORIEFF, N., MUSACCHIO, A. & NOGALES, E. 2010. The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. *Nature,* 467**,** 805-10.

AN, S., KOLDEWEY, P., CHIK, J., SUBRAMANIAN, L. & CHO, U. S. 2018. Mis16 Switches Function from a Histone H4 Chaperone to a CENP-A(Cnp1)-Specific Assembly Factor through Eic1 Interaction. *Structure*.

ANDERSON, M. H., J.; YEH, E.; BLOOM, K. 2009. Function and Assembly of DNA Looping, Clustering, and Microtubule Attachment Complexes within a Eukaryotic Kinetochore. *Mol Biol Cell,* 20**,** 4131-4139.

ARAVAMUDHAN, P., GOLDFARB, A. A. & JOGLEKAR, A. P. 2015. The kinetochore encodes a mechanical switch to disrupt spindle assembly checkpoint signalling. *Nat Cell Biol*.

ARAVIND, L., IYER, L. M. & WU, C. 2007. Domain architectures of the Scm3p protein provide insights into centromere function and evolution. *Cell Cycle,* 6**,** 2511-5.

ASBURY, C. L., GESTAUT, D. R., POWERS, A. F., FRANCK, A. D. & DAVIS, T. N. 2006. The Dam1 kinetochore complex harnesses microtubule dynamics to produce force and movement. *Proc Natl Acad Sci U S A,* 103**,** 9873-8.

BAI, X. C., MCMULLAN, G. & SCHERES, S. H. 2015. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci,* 40**,** 49-57.

BAKER, L. A. & RUBINSTEIN, J. L. 2010. Radiation damage in electron cryomicroscopy. *Methods Enzymol,* 481**,** 371-88.

BANSAL, P. K., ABDULLE, R. & KITAGAWA, K. 2004. Sgt1 associates with Hsp90: an initial step of assembly of the core kinetochore complex. *Mol Cell Biol,* 24**,** 8069-79.

BARTESAGHI, A., AGUERREBERE, C., FALCONIERI, V., BANERJEE, S., EARL, L. A., ZHU, X., GRIGORIEFF, N., MILNE, J. L. S., SAPIRO, G., WU, X. & SUBRAMANIAM, S. 2018. Atomic Resolution Cryo-EM Structure of beta-Galactosidase. *Structure*.

BELLIZZI, J. J. S., P.K.; HARRISON, S.C. 2007. Crystal Structure of the Yeast Inner Kinetochore Subunit Cep3p. *Structure,* 15**,** 1422-1430.

BELTRAO, P., ALBANESE, V., KENNER, L. R., SWANEY, D. L., BURLINGAME, A., VILLEN, J., LIM, W. A., FRASER, J. S., FRYDMAN, J. & KROGAN, N. J. 2012.

Systematic functional prioritization of protein posttranslational modifications. *Cell,* 150**,** 413-25.
BERNAT, R. L., DELANNOY, M. R., ROTHFIELD, N. F. & EARNSHAW, W. C. 1991. Disruption of centromere assembly during interphase inhibits kinetochore morphogenesis and function in mitosis. *Cell,* 66**,** 1229-38.
BIGGINS, S. 2013. The composition, functions, and regulation of the budding yeast kinetochore. *Genetics,* 194**,** 817-46.
BLACK, B. E., FOLTZ, D. R., CHAKRAVARTHY, S., LUGER, K., WOODS, V. L., JR. & CLEVELAND, D. W. 2004. Structural determinants for generating centromeric chromatin. *Nature,* 430**,** 578-82.
BODOR, D. L., MATA, J. F., SERGEEV, M., DAVID, A. F., SALIMIAN, K. J., PANCHENKO, T., CLEVELAND, D. W., BLACK, B. E., SHAH, J. V. & JANSEN, L. E. 2014. The quantitative architecture of centromeric chromatin. *Elife,* 3**,** e02137.
BOLAND, A., CHANG, L. & BARFORD, D. 2017. The potential of cryo-electron microscopy for structure-based drug design. *Essays Biochem,* 61**,** 543-560.
BOND, S. R. & NAUS, C. C. 2012. RF-Cloning.org: an online tool for the design of restriction-free cloning projects. *Nucleic Acids Res,* 40**,** W209-13.
BOUCK, D. C. & BLOOM, K. S. 2005. The kinetochore protein Ndc10p is required for spindle stability and cytokinesis in yeast. *Proc Natl Acad Sci U S A,* 102**,** 5408-13.
BRAM, R. J. & KORNBERG, R. D. 1987. Isolation of a Saccharomyces cerevisiae centromere DNA-binding protein, its human homolog, and its possible role as a transcription factor. *Mol Cell Biol,* 7**,** 403-9.
BRIGGS, J. A. 2013. Structural biology in situ--the potential of subtomogram averaging. *Curr Opin Struct Biol,* 23**,** 261-7.
BRILOT, A. F., CHEN, J. Z., CHENG, A., PAN, J., HARRISON, S. C., POTTER, C. S., CARRAGHER, B., HENDERSON, R. & GRIGORIEFF, N. 2012. Beam-induced motion of vitrified specimen on holey carbon film. *J Struct Biol,* 177**,** 630-7.
BROWN, M. T., GOETSCH, L. & HARTWELL, L. H. 1993. MIF2 is required for mitotic spindle integrity during anaphase spindle elongation in Saccharomyces cerevisiae. *J Cell Biol,* 123**,** 387-403.
BRÜGGELLER, P. & MAYER, E. 1980. Complete vitrification in pure liquid water and dilute aqueous solutions. *Nature,* 288**,** 569-571.
BURRACK, L. S. & BERMAN, J. 2012. Flexibility of centromere and kinetochore structures. *Trends Genet,* 28**,** 204-12.
CAI, M. J. & DAVIS, R. W. 1989. Purification of a yeast centromere-binding protein that is able to distinguish single base-pair mutations in its recognition site. *Mol Cell Biol,* 9**,** 2544-50.
CAMAHORT, R., LI, B., FLORENS, L., SWANSON, S. K., WASHBURN, M. P. & GERTON, J. L. 2007. Scm3 is essential to recruit the histone h3 variant cse4 to centromeres and to maintain a functional kinetochore. *Mol Cell,* 26**,** 853-65.
CAMAHORT, R., SHIVARAJU, M., MATTINGLY, M., LI, B., NAKANISHI, S., ZHU, D., SHILATIFARD, A., WORKMAN, J. L. & GERTON, J. L. 2009. Cse4 is part of an octameric nucleosome in budding yeast. *Mol Cell,* 35**,** 794-805.
CAMPBELL, M. G., CHENG, A., BRILOT, A. F., MOELLER, A., LYUMKIS, D., VEESLER, D., PAN, J., HARRISON, S. C., POTTER, C. S., CARRAGHER, B. & GRIGORIEFF, N. 2012. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure,* 20**,** 1823-8.
CANE, S., YE, A. A., LUKS-MORGAN, S. J. & MARESCA, T. J. 2013. Elevated polar ejection forces stabilize kinetochore-microtubule attachments. *J Cell Biol,* 200**,** 203-18.

CAREY, M. F., PETERSON, C. L. & SMALE, S. T. 2013. PCR-mediated site-directed mutagenesis. *Cold Spring Harb Protoc,* 2013**,** 738-42.
CARMENA, M., RUCHAUD, S. & EARNSHAW, W. C. 2009. Making the Auroras glow: regulation of Aurora A and B kinase function by interacting proteins. *Curr Opin Cell Biol,* 21**,** 796-805.
CARROLL, C. W., MILKS, K. J. & STRAIGHT, A. F. 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J Cell Biol,* 189**,** 1143-55.
CARROLL, C. W., SILVA, M. C., GODEK, K. M., JANSEN, L. E. & STRAIGHT, A. F. 2009. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nat Cell Biol,* 11**,** 896-902.
CHAN, Y. W., JEYAPRAKASH, A. A., NIGG, E. A. & SANTAMARIA, A. 2012. Aurora B controls kinetochore-microtubule attachments by inhibiting Ska complex-KMN network interaction. *J Cell Biol,* 196**,** 563-71.
CHEERAMBATHUR, D. K. & DESAI, A. 2014. Linked in: formation and regulation of microtubule attachments during chromosome segregation. *Curr Opin Cell Biol,* 26**,** 113-22.
CHEESEMAN, I. M., CHAPPIE, J. S., WILSON-KUBALEK, E. M. & DESAI, A. 2006. The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell,* 127**,** 983-97.
CHEN, J. & LIU, J. 2014. Spatial-temporal model for silencing of the mitotic spindle assembly checkpoint. *Nat Commun,* 5**,** 4795.
CHEN, Y., BAKER, R. E., KEITH, K. C., HARRIS, K., STOLER, S. & FITZGERALD-HAYES, M. 2000. The N terminus of the centromere H3-like protein Cse4p performs an essential function distinct from that of the histone fold domain. *Mol Cell Biol,* 20**,** 7037-48.
CHENG, Y., GRIGORIEFF, N., PENCZEK, P. A. & WALZ, T. 2015. A Primer to Single-Particle Cryo-Electron Microscopy. *Cell,* 161**,** 438-449.
CHO, U. S. & HARRISON, S. C. 2011. Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. *Proc Natl Acad Sci U S A,* 108**,** 9367-71.
CHO, U. S. & HARRISON, S. C. 2012. Ndc10 is a platform for inner kinetochore assembly in budding yeast. *Nat Struct Mol Biol,* 19**,** 48-55.
CIFERRI, C., PASQUALATO, S., SCREPANTI, E., VARETTI, G., SANTAGUIDA, S., DOS REIS, G., MAIOLICA, A., POLKA, J., DE LUCA, J. G., DE WULF, P., SALEK, M., RAPPSILBER, J., MOORES, C. A., SALMON, E. D. & MUSACCHIO, A. 2008. Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. *Cell,* 133**,** 427-39.
CLARKE, L. & CARBON, J. 1980. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature,* 287**,** 504-9.
COHEN, R. L., ESPELIN, C. W., DE WULF, P., SORGER, P. K., HARRISON, S. C. & SIMONS, K. T. 2008. Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. *Mol Biol Cell,* 19**,** 4480-91.
COLE, H. A., HOWARD, B. H. & CLARK, D. J. 2011. The centromeric nucleosome of budding yeast is perfectly positioned and covers the entire centromere. *Proc Natl Acad Sci U S A,* 108**,** 12687-92.
CONNELLY, C. & HIETER, P. 1996. Budding Yeast SKP1 Encodes an Evolutionarily Conserved Kinetochore Protein Required for Cell Cycle Progression. *Cell,* 86**,** 275-285.
DALAL, Y. W., H.; LINDSAY, S.; HENIKOFF, S. 2007. Tetrameric Structure of Centromeric Nucleosomes in Interphase Drosophila Cells. *PLoS Biology,* 5.
DE CARLO, S. & HARRIS, J. R. 2011. Negative staining and cryo-negative staining of macromolecules and viruses for TEM. *Micron,* 42**,** 117-31.
DE LA ROSA-TREVIN, J. M., QUINTANA, A., DEL CANO, L., ZALDIVAR, A., FOCHE, I., GUTIERREZ, J., GOMEZ-BLANCO, J., BURGUET-CASTELL, J., CUENCA-

ALBA, J., ABRISHAMI, V., VARGAS, J., OTON, J., SHAROV, G., VILAS, J. L., NAVAS, J., CONESA, P., KAZEMI, M., MARABINI, R., SORZANO, C. O. & CARAZO, J. M. 2016. Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J Struct Biol*.

DE WULF, P., MCAINSH, A. D. & SORGER, P. K. 2003. Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. *Genes Dev,* 17**,** 2902-21.

DECHASSA, M. L., WYNS, K. & LUGER, K. 2014. Scm3 deposits a (Cse4-H4)2 tetramer onto DNA through a Cse4-H4 dimer intermediate. *Nucleic Acids Res,* 42**,** 5532-42.

DELUCA, J. G., GALL, W. E., CIFERRI, C., CIMINI, D., MUSACCHIO, A. & SALMON, E. D. 2006. Kinetochore microtubule dynamics and attachment stability are regulated by Hec1. *Cell,* 127**,** 969-82.

DELUCA, J. G., HOWELL, B. J., CANMAN, J. C., HICKEY, J. M., FANG, G. & SALMON, E. D. 2003. Nuf2 and Hec1 are required for retention of the checkpoint proteins Mad1 and Mad2 to kinetochores. *Curr Biol,* 13**,** 2103-9.

DELUCA, J. G. & MUSACCHIO, A. 2012. Structural organization of the kinetochore-microtubule interface. *Curr Opin Cell Biol,* 24**,** 48-56.

DESHAIES, R. J. 1999. SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu Rev Cell Dev Biol,* 15**,** 435-67.

DIAZ-INGELMO, O., MARTINEZ-GARCIA, B., SEGURA, J., VALDES, A. & ROCA, J. 2015. DNA Topology and Global Architecture of Point Centromeres. *Cell Rep,* 13**,** 667-77.

DIMITROVA, Y. N., JENNI, S., VALVERDE, R., KHIN, Y. & HARRISON, S. C. 2016. Structure of the MIND Complex Defines a Regulatory Focus for Yeast Kinetochore Assembly. *Cell,* 167**,** 1014-1027 e12.

DION, V., KALCK, V., HORIGOME, C., TOWBIN, B. D. & GASSER, S. M. 2012. Increased mobility of double-strand breaks requires Mec1, Rad9 and the homologous recombination machinery. *Nat Cell Biol,* 14**,** 502-9.

DOHENY, K. F., SORGER, P. K., HYMAN, A. A., TUGENDREICH, S., SPENCER, F. & HIETER, P. 1993. Identification of essential components of the S. cerevisiae kinetochore. *Cell,* 73**,** 761-74.

DOU, Z., LIU, X., WANG, W., ZHU, T., WANG, X., XU, L., ABRIEU, A., FU, C., HILL, D. L. & YAO, X. 2015. Dynamic localization of Mps1 kinase to kinetochores is essential for accurate spindle microtubule attachment. *Proc Natl Acad Sci U S A,* 112**,** E4546-55.

DUBOCHET, J., ADRIAN, M., CHANG, J. J., HOMO, J. C., LEPAULT, J., MCDOWALL, A. W. & SCHULTZ, P. 1988. Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys,* 21**,** 129-228.

DUBOCHET, J. & BOOY, F. P., R FREEMAN, A V JONES, AND C A WALTERJ DUBOCHET, F P BOOY, R FREEMAN, A V JONES, AND C A WALTER 1981. Low Temperature Electron Microscopy. *Ann.Rev.Biophys. Bioeng.,* 10**,** 133-49.

DUBOCHET, J. & MCDOWALL, A. W. 1981. VITRIFICATION OF PURE WATER FOR ELECTRON MICROSCOPY. *Journal of Microscopy,* 124**,** 3-4.

EARNSHAW, W. C. & ROTHFIELD, N. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma,* 91**,** 313-21.

EGELMAN, E. H. 2016. The Current Revolution in Cryo-EM. *Biophys J,* 110**,** 1008-12.

EMANUELE, M. J., LAN, W., JWA, M., MILLER, S. A., CHAN, C. S. & STUKENBERG, P. T. 2008. Aurora B kinase and protein phosphatase 1 have opposing roles in modulating kinetochore assembly. *J Cell Biol,* 181**,** 241-54.

EMSLEY, P. & COWTAN, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr,* 60**,** 2126-32.

ENQUIST-NEWMAN, M., CHEESEMAN, I. M., VAN GOOR, D., DRUBIN, D. G., MELUH, P. B. & BARNES, G. 2001. Dad1p, third component of the Duo1p/Dam1p complex involved in kinetochore function and mitotic spindle integrity. *Mol Biol Cell,* 12**,** 2601-13.

ESPELIN, C. W. K., K.B.; SORGER, P.K. 1997. Probing the Architecture of a Simple Kinetochore Using DNA–Protein Crosslinking. *J Cell Biol,* 139**,** 1383-1396.

ESPELIN, C. W. S., K.T.; HARRISON, S.C.; SORGER, P.K. 2003. Binding of the Essential Saccharomyces cerevisiae Kinetochore Protein Ndc10p to CDEII. *Mol Biol Cell,* 14**,** 4557-4568.

ESPERT, A., ULUOCAK, P., BASTOS, R. N., MANGAT, D., GRAAB, P. & GRUNEBERG, U. 2014. PP2A-B56 opposes Mps1 phosphorylation of Knl1 and thereby promotes spindle assembly checkpoint silencing. *J Cell Biol,* 206**,** 833-42.

EUSKIRCHEN, G. M. 2002. Nnf1p, Dsn1p, Mtw1p, and Nsl1p: a new group of proteins important for chromosome segregation in Saccharomyces cerevisiae. *Eukaryot Cell,* 1**,** 229-40.

FALK, S. J., GUO, L. Y., SEKULIC, N., SMOAK, E. M., MANI, T., LOGSDON, G. A., GUPTA, K., JANSEN, L. E., VAN DUYNE, G. D., VINOGRADOV, S. A., LAMPSON, M. A. & BLACK, B. E. 2015. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science,* 348**,** 699-703.

FISCHER, N., NEUMANN, P., KONEVEGA, A. L., BOCK, L. V., FICNER, R., RODNINA, M. V. & STARK, H. 2015. Structure of the E. coli ribosome-EF-Tu complex at <3 A resolution by Cs-corrected cryo-EM. *Nature,* 520**,** 567-70.

FITZGERALD, D. J., BERGER, P., SCHAFFITZEL, C., YAMADA, K., RICHMOND, T. J. & BERGER, I. 2006. Protein complex expression by using multigene baculoviral vectors. *Nat Methods,* 3**,** 1021-32.

FOLTZ, D. R., JANSEN, L. E., BAILEY, A. O., YATES, J. R., 3RD, BASSETT, E. A., WOOD, S., BLACK, B. E. & CLEVELAND, D. W. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell,* 137**,** 472-84.

FRANCISCO, L., WANG, W. & CHAN, C. S. 1994. Type 1 protein phosphatase acts in opposition to IpL1 protein kinase in regulating yeast chromosome segregation. *Mol Cell Biol,* 14**,** 4731-40.

FRANK, J. 2006. *Three-dimensional electron microscopy of macromolecular assemblies : visualization of biological molecules in their native state,* Oxford ; New York, Oxford University Press.

FUJITA, Y., HAYASHI, T., KIYOMITSU, T., TOYODA, Y., KOKUBU, A., OBUSE, C. & YANAGIDA, M. 2007. Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev Cell,* 12**,** 17-30.

FURTH, N., GERTMAN, O., SHIBER, A., ALFASSY, O. S., COHEN, I., ROSENBERG, M. M., DORON, N. K., FRIEDLER, A. & RAVID, T. 2011. Exposure of bipartite hydrophobic signal triggers nuclear quality control of Ndc10 at the endoplasmic reticulum/nuclear envelope. *Mol Biol Cell,* 22**,** 4726-39.

FURUYAMA, S. & BIGGINS, S. 2007. Centromere identity is specified by a single centromeric nucleosome in budding yeast. *Proc Natl Acad Sci U S A,* 104**,** 14706-11.

FURUYAMA, T., CODOMO, C. A. & HENIKOFF, S. 2013. Reconstitution of hemisomes on budding yeast centromeric DNA. *Nucleic Acids Res,* 41**,** 5769-83.

FURUYAMA, T. & HENIKOFF, S. 2009. Centromeric nucleosomes induce positive DNA supercoils. *Cell,* 138**,** 104-13.

GAITANOS, T. N., SANTAMARIA, A., JEYAPRAKASH, A. A., WANG, B., CONTI, E. & NIGG, E. A. 2009. Stable kinetochore-microtubule interactions depend on the Ska complex and its new component Ska3/C13Orf3. *EMBO J,* 28**,** 1442-52.

GARDNER, R. D. P., A.; YELLMAN, C.; TAVORMINA, P.A.; MONTEAGUDO, M.C.; BURKE, D.J. 2001. The Spindle Checkpoint of the Yeast Saccharomyces cerevisiae Requires Kinetochore Function and Maps to the CBF3 Domain. *Genetics,* 157**,** 1493-1502.

GIBSON, D. G., YOUNG, L., CHUANG, R. Y., VENTER, J. C., HUTCHISON, C. A., 3RD & SMITH, H. O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods,* 6**,** 343-5.

GILLIS, A. N., THOMAS, S., HANSEN, S. D. & KAPLAN, K. B. 2005. A novel role for the CBF3 kinetochore-scaffold complex in regulating septin dynamics and cytokinesis. *J Cell Biol,* 171**,** 773-84.

GLAESER, R. M. & HAN, B. G. 2017. Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophys Rep,* 3**,** 1-7.

GOH, P. K., J.V. 1993. NDC10: A Gene Involved in Chromosome Segregation in Saccharomyces cerevisiae. *J Cell Biol,* 121**,** 503-512.

HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.

HANISCH, A., SILLJE, H. H. & NIGG, E. A. 2006. Timely anaphase onset requires a novel spindle and kinetochore complex comprising Ska1 and Ska2. *EMBO J,* 25**,** 5504-15.

HANSSON, M. D., RZEZNICKA, K., ROSENBACK, M., HANSSON, M. & SIRIJOVSKI, N. 2008. PCR-mediated deletion of plasmid DNA. *Anal Biochem,* 375**,** 373-5.

HAO, B., ZHENG, N., SCHULMAN, B. A., WU, G., MILLER, J. J., PAGANO, M. & PAVLETICH, N. P. 2005. Structural basis of the Cks1-dependent recognition of p27(Kip1) by the SCF(Skp2) ubiquitin ligase. *Mol Cell,* 20**,** 9-19.

HARDWICK, K. G., JOHNSTON, R. C., SMITH, D. L. & MURRAY, A. W. 2000. MAD3 encodes a novel component of the spindle checkpoint which interacts with Bub3p, Cdc20p, and Mad2p. *J Cell Biol,* 148**,** 871-82.

HAYASHI, T., FUJITA, Y., IWASAKI, O., ADACHI, Y., TAKAHASHI, K. & YANAGIDA, M. 2004. Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell,* 118**,** 715-29.

HE, X., RINES, D. R., ESPELIN, C. W. & SORGER, P. K. 2001. Molecular analysis of kinetochore-microtubule attachment in budding yeast. *Cell,* 106**,** 195-206.

HEGEMANN, J. H. & FLEIG, U. N. 1993. The centromere of budding yeast. *Bioessays,* 15**,** 451-60.

HEGEMANN, J. H., SHERO, J. H., COTTAREL, G., PHILIPPSEN, P. & HIETER, P. 1988. Mutational analysis of centromere DNA from chromosome VI of Saccharomyces cerevisiae. *Mol Cell Biol,* 8**,** 2523-35.

HENDERSON, R. & MCMULLAN, G. 2013. Problems in obtaining perfect images by single-particle electron cryomicroscopy of biological structures in amorphous ice. *Microscopy (Oxf),* 62**,** 43-50.

HIRUMA, Y., SACRISTAN, C., PACHIS, S. T., ADAMOPOULOS, A., KUIJT, T., UBBINK, M., VON CASTELMUR, E., PERRAKIS, A. & KOPS, G. J. 2015. Competition between MPS1 and microtubules at kinetochores regulates spindle checkpoint signaling. *Science,* 348**,** 1264-7.

HOFMANN, C., CHEESEMAN, I. M., GOODE, B. L., MCDONALD, K. L., BARNES, G. & DRUBIN, D. G. 1998. Saccharomyces cerevisiae Duo1p and Dam1p, novel proteins involved in mitotic spindle function. *J Cell Biol,* 143**,** 1029-40.

HOLT, L. J., TUCH, B. B., VILLEN, J., JOHNSON, A. D., GYGI, S. P. & MORGAN, D. O. 2009. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science,* 325**,** 1682-6.

HORI, T., AMANO, M., SUZUKI, A., BACKER, C. B., WELBURN, J. P., DONG, Y., MCEWEN, B. F., SHANG, W. H., SUZUKI, E., OKAWA, K., CHEESEMAN, I. M. & FUKAGAWA, T. 2008. CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell,* 135**,** 1039-52.

HORNUNG, P., MAIER, M., ALUSHIN, G. M., LANDER, G. C., NOGALES, E. & WESTERMANN, S. 2011. Molecular architecture and connectivity of the budding yeast Mtw1 kinetochore complex. *J Mol Biol,* 405**,** 548-59.

HORNUNG, P., TROC, P., MALVEZZI, F., MAIER, M., DEMIANOVA, Z., ZIMNIAK, T., LITOS, G., LAMPERT, F., SCHLEIFFER, A., BRUNNER, M., MECHTLER, K., HERZOG, F., MARLOVITS, T. C. & WESTERMANN, S. 2014. A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. *J Cell Biol,* 206**,** 509-24.

HOWELL, B. J., MCEWEN, B. F., CANMAN, J. C., HOFFMAN, D. B., FARRAR, E. M., RIEDER, C. L. & SALMON, E. D. 2001. Cytoplasmic dynein/dynactin drives kinetochore protein transport to the spindle poles and has a role in mitotic spindle checkpoint inactivation. *J Cell Biol,* 155**,** 1159-72.

HU, H., LIU, Y., WANG, M., FANG, J., HUANG, H., YANG, N., LI, Y., WANG, J., YAO, X., SHI, Y., LI, G. & XU, R. M. 2011. Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. *Genes Dev,* 25**,** 901-6.

HUIS IN 'T VELD, P. J., JEGANATHAN, S., PETROVIC, A., SINGH, P., JOHN, J., KRENN, V., WEISSMANN, F., BANGE, T. & MUSACCHIO, A. 2016. Molecular basis of outer kinetochore assembly on CENP-T. *Elife,* 5.

HWANG, L. H., LAU, L. F., SMITH, D. L., MISTROT, C. A., HARDWICK, K. G., HWANG, E. S., AMON, A. & MURRAY, A. W. 1998. Budding yeast Cdc20: a target of the spindle checkpoint. *Science,* 279**,** 1041-4.

HYLAND, K. M., KINGSBURY, J., KOSHLAND, D. & HIETER, P. 1999. Ctf19p: A novel kinetochore protein in Saccharomyces cerevisiae and a potential link between the kinetochore and mitotic spindle. *J Cell Biol,* 145**,** 15-28.

IRNIGER, S., PIATTI, S., MICHAELIS, C. & NASMYTH, K. 1995. Genes involved in sister chromatid separation are needed for B-type cyclin proteolysis in budding yeast. *Cell,* 81**,** 269-78.

IZAWA, D. & PINES, J. 2015. The mitotic checkpoint complex binds a second CDC20 to inhibit active APC/C. *Nature,* 517**,** 631-4.

JALLEPALLI, P. V. & LENGAUER, C. 2001. Chromosome segregation and cancer: cutting through the mystery. *Nat Rev Cancer,* 1**,** 109-17.

JANKE, C., ORTIZ, J., LECHNER, J., SHEVCHENKO, A., SHEVCHENKO, A., MAGIERA, M. M., SCHRAMM, C. & SCHIEBEL, E. 2001. The budding yeast proteins Spc24p and Spc25p interact with Ndc80p and Nuf2p at the kinetochore and are important for kinetochore clustering and checkpoint control. *EMBO J,* 20**,** 777-91.

JEHN, B. N., R.; HEGEMANN, J.H. 1991. In Vivo Analysis of the Saccharomyces cerevisiae Centromere CDEIII Sequence: Requirements for Mitotic Chromosome Segregation. *Mol Cell Biol,* 11**,** 5212-5221.

JI, Z., GAO, H. & YU, H. 2015. Kinetochore attachment sensed by competitive Mps1 and microtubule binding to Ndc80C. *Science,* 348**,** 1260-4.

JIA, L., LI, B. & YU, H. 2016. The Bub1-Plk1 kinase complex promotes spindle checkpoint signalling through Cdc20 phosphorylation. *Nat Commun,* 7**,** 10818.

JIANG, W. L., J.; CARBON, J. 1993. Isolation and Characterization of a Gene (CBF2)Specifying
a Protein Component of the Budding Yeast Kinetochore. *J Cell Biol,* 121**,** 513-519.

JOGLEKAR, A. P. 2016. A Cell Biological Perspective on Past, Present and Future Investigations of the Spindle Assembly Checkpoint. *Biology (Basel),* 5.

JOGLEKAR, A. P., BLOOM, K. & SALMON, E. D. 2009. In vivo protein architecture of the eukaryotic kinetochore with nanometer scale accuracy. *Curr Biol,* 19**,** 694-9.

JOGLEKAR, A. P., BLOOM, K. S. & SALMON, E. D. 2010. Mechanisms of force generation by end-on kinetochore-microtubule attachments. *Curr Opin Cell Biol,* 22**,** 57-67.

JOGLEKAR, A. P., BOUCK, D., FINLEY, K., LIU, X., WAN, Y., BERMAN, J., HE, X., SALMON, E. D. & BLOOM, K. S. 2008. Molecular architecture of the kinetochore-microtubule attachment site is conserved between point and regional centromeres. *J Cell Biol,* 181**,** 587-94.

JONES, M. H., HE, X., GIDDINGS, T. H. & WINEY, M. 2001. Yeast Dam1p has a role at the kinetochore in assembly of the mitotic spindle. *Proc Natl Acad Sci U S A,* 98**,** 13675-80.

JONIC, S., SORZANO, C. O. & BOISSET, N. 2008. Comparison of single-particle analysis and electron tomography approaches: an overview. *J Microsc,* 232**,** 562-79.

KAITNA, S., MENDOZA, M., JANTSCH-PLUNGER, V. & GLOTZER, M. 2000. Incenp and an aurora-like kinase form a complex essential for chromosome segregation and efficient completion of cytokinesis. *Curr Biol,* 10**,** 1172-81.

KAO, L., WANG, Y. T., CHEN, Y. C., TSENG, S. F., JHANG, J. C., CHEN, Y. J. & TENG, S. C. 2014. Global analysis of cdc14 dephosphorylation sites reveals essential regulatory role in mitosis and cytokinesis. *Mol Cell Proteomics,* 13**,** 594-605.

KAPLAN, K. B. H., A.A.; SORGER, P.K. 1997. Regulating the Yeast Kinetochore by Ubiquitin-Dependent Degradation and Skp1p-Mediated Phosphorylation. *Cell,* 91**,** 491-500.

KATO, H., JIANG, J., ZHOU, B. R., ROZENDAAL, M., FENG, H., GHIRLANDO, R., XIAO, T. S., STRAIGHT, A. F. & BAI, Y. 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science,* 340**,** 1110-3.

KAWASHIMA, S. A., YAMAGISHI, Y., HONDA, T., ISHIGURO, K. & WATANABE, Y. 2010. Phosphorylation of H2A by Bub1 prevents chromosomal instability through localizing shugoshin. *Science,* 327**,** 172-7.

KEATING, P., RACHIDI, N., TANAKA, T. U. & STARK, M. J. 2009. Ipl1-dependent phosphorylation of Dam1 is reduced by tension applied on kinetochores. *J Cell Sci,* 122**,** 4375-82.

KEITH, K. C., BAKER, R. E., CHEN, Y., HARRIS, K., STOLER, S. & FITZGERALD-HAYES, M. 1999. Analysis of primary structural determinants that distinguish the centromere-specific function of histone variant Cse4p from histone H3. *Mol Cell Biol,* 19**,** 6130-9.

KELLY, A. E., GHENOIU, C., XUE, J. Z., ZIERHUT, C., KIMURA, H. & FUNABIKI, H. 2010. Survivin reads phosphorylated histone H3 threonine 3 to activate the mitotic kinase Aurora B. *Science,* 330**,** 235-9.

KEMMLER, S., STACH, M., KNAPP, M., ORTIZ, J., PFANNSTIEL, J., RUPPERT, T. & LECHNER, J. 2009. Mimicking Ndc80 phosphorylation triggers spindle assembly checkpoint signalling. *EMBO J,* 28**,** 1099-110.

KHOSHOUEI, M., RADJAINIA, M., BAUMEISTER, W. & DANEV, R. 2017. Cryo-EM structure of haemoglobin at 3.2 A determined with the Volta phase plate. *Nat Commun,* 8**,** 16099.

KING, D. A., ZHANG, L., GUARENTE, L. & MARMORSTEIN, R. 1999. Structure of a HAP1-DNA complex reveals dramatically asymmetric DNA binding by a homodimeric protein. *Nat Struct Biol,* 6**,** 64-71.

KINGSTON, I. J., YUNG, J. S. & SINGLETON, M. R. 2011. Biophysical characterization of the centromere-specific nucleosome from budding yeast. *J Biol Chem,* 286**,** 4021-6.

KIPREOS, E. T. & PAGANO, M. 2000. The F-box protein family. *Genome Biol,* 1**,** REVIEWS3002.

KISELEV, N. A., SHERMAN, M. B. & TSUPRUN, V. L. 1990. Negative staining of proteins. *Electron Microsc Rev,* 3**,** 43-72.

KITAGAWA, K., ABDULLE, R., BANSAL, P. K., CAGNEY, G., FIELDS, S. & HIETER, P. 2003. Requirement of Skp1-Bub1 interaction for kinetochore-mediated activation of the spindle checkpoint. *Mol Cell,* 11**,** 1201-13.

KITAGAWA, K., SKOWYRA, D., ELLEDGE, S. J., HARPER, J. W. & HIETER, P. 1999. SGT1 Encodes an Essential Component of the Yeast Kinetochore Assembly Pathway and a Novel Subunit of the SCF Ubiquitin Ligase Complex. *Molecular Cell,* 4**,** 21-33.

KRAWCZYK, P. M., BOROVSKI, T., STAP, J., CIJSOUW, T., TEN CATE, R., MEDEMA, J. P., KANAAR, R., FRANKEN, N. A. & ATEN, J. A. 2012. Chromatin mobility is increased at sites of DNA double-strand breaks. *J Cell Sci,* 125**,** 2127-33.

KRENN, V. & MUSACCHIO, A. 2015. The Aurora B Kinase in Chromosome Bi-Orientation and Spindle Checkpoint Signaling. *Front Oncol,* 5**,** 225.

KUHLBRANDT, W. 2014. Cryo-EM enters a new era. *Elife,* 3**,** e03678.

KUILMAN, T., MAIOLICA, A., GODFREY, M., SCHEIDEL, N., AEBERSOLD, R. & UHLMANN, F. 2015. Identification of Cdk targets that control cytokinesis. *EMBO J,* 34**,** 81-96.

KUMANOMIDOU, T., NISHIO, K., TAKAGI, K., NAKAGAWA, T., SUZUKI, A., YAMANE, T., TOKUNAGA, F., IWAI, K., MURAKAMI, A., YOSHIDA, Y., TANAKA, K. & MIZUSHIMA, T. 2015. The Structural Differences between a Glycoprotein Specific F-Box Protein Fbs1 and Its Homologous Protein FBG3. *PLoS One,* 10**,** e0140366.

LACEFIELD, S. 2017. Chromosome Biology: Specification of the Kinetochore for Cohesin Recruitment. *Curr Biol,* 27**,** R1319-R1321.

LAMPERT, F., HORNUNG, P. & WESTERMANN, S. 2010. The Dam1 complex confers microtubule plus end-tracking activity to the Ndc80 kinetochore complex. *J Cell Biol,* 189**,** 641-9.

LAMPSON, M. A. & GRISHCHUK, E. L. 2017. Mechanisms to Avoid and Correct Erroneous Kinetochore-Microtubule Attachments. *Biology (Basel),* 6.

LARA-GONZALEZ, P., WESTHORPE, F. G. & TAYLOR, S. S. 2012. The spindle assembly checkpoint. *Curr Biol,* 22**,** R966-80.

LAWRIMORE, J., VASQUEZ, P. A., FALVO, M. R., TAYLOR, R. M., 2ND, VICCI, L., YEH, E., FOREST, M. G. & BLOOM, K. 2015. DNA loops generate intracentromere tension in mitosis. *J Cell Biol,* 210**,** 553-64.

LEBER, V., NANS, A. & SINGLETON, M.R. 2018. Structural basis for the assembly of the CBF3 kinetochore complex. *EMBO J*, 37, 269-281.

LECHNER, J. 1994. A zinc finger protein, essential for chromosome segregation, constitutes a putative DNA binding subunit of the Saccharomyces cerevisiae kinetochore complex, Cbf3. *EMBO J,* 13**,** 5203-11.

LECHNER, J. C., J. 1991. A 240 kd Multisubunit Protein Complex, CBF3, Is a Major Component of the Budding Yeast Centromere. *Cell,* 64**,** 717-725.

LI, X., MOONEY, P., ZHENG, S., BOOTH, C. R., BRAUNFELD, M. B., GUBBENS, S., AGARD, D. A. & CHENG, Y. 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods,* 10**,** 584-90.

LI, Y. & HAO, B. 2010. Structural basis of dimerization-dependent ubiquitination by the SCF(Fbx4) ubiquitin ligase. *J Biol Chem,* 285**,** 13896-906.

LIANG, S. D., MARMORSTEIN, R., HARRISON, S. C. & PTASHNE, M. 1996. DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA. *Mol Cell Biol,* 16**,** 3773-80.

LIAO, M., CAO, E., JULIUS, D. & CHENG, Y. 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature,* 504**,** 107-12.

LINGELBACH, L. B. & KAPLAN, K. B. 2004. The interaction between Sgt1p and Skp1p is regulated by HSP90 chaperones and is required for proper CBF3 assembly. *Mol Cell Biol,* 24**,** 8938-50.

LIU, D., VADER, G., VROMANS, M. J., LAMPSON, M. A. & LENS, S. M. 2009. Sensing chromosome bi-orientation by spatial separation of aurora B kinase from kinetochore substrates. *Science,* 323**,** 1350-3.

LIU, D., VLEUGEL, M., BACKER, C. B., HORI, T., FUKAGAWA, T., CHEESEMAN, I. M. & LAMPSON, M. A. 2010. Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. *J Cell Biol,* 188**,** 809-20.

LIU, S. T., HITTLE, J. C., JABLONSKI, S. A., CAMPBELL, M. S., YODA, K. Y. & EN, T. J. 2003. Human CENP-I specifies localization of CENP-F, MAD1 and MAD2 to kinetochores and is essential for mitosis. *Nat Cell Biol,* 5**,** 341-345.

LONDON, N., CETO, S., RANISH, J. A. & BIGGINS, S. 2012. Phosphoregulation of Spc105 by Mps1 and PP1 regulates Bub1 localization to kinetochores. *Curr Biol,* 22**,** 900-6.

LOTTERSBERGER, F., KARSSEMEIJER, R. A., DIMITROVA, N. & DE LANGE, T. 2015. 53BP1 and the LINC Complex Promote Microtubule-Dependent DSB Mobility and DNA Repair. *Cell,* 163**,** 880-93.

LOWARY, P. T. & WIDOM, J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol,* 276**,** 19-42.

MAKRANTONI, V. & STARK, M. J. 2009. Efficient chromosome biorientation and the tension checkpoint in Saccharomyces cerevisiae both require Bir1. *Mol Cell Biol,* 29**,** 4552-62.

MALIK, H. S. & HENIKOFF, S. 2002. Conflict begets complexity: the evolution of centromeres. *Curr Opin Genet Dev,* 12**,** 711-8.

MALIK, H. S. & HENIKOFF, S. 2009. Major evolutionary transitions in centromere complexity. *Cell,* 138**,** 1067-82.

MALVEZZI, F., LITOS, G., SCHLEIFFER, A., HEUCK, A., MECHTLER, K., CLAUSEN, T. & WESTERMANN, S. 2013. A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. *EMBO J,* 32**,** 409-23.

MARMORSTEIN, R., CAREY, M., PTASHNE, M. & HARRISON, S. C. 1992. DNA recognition by GAL4: structure of a protein-DNA complex. *Nature,* 356**,** 408-14.

MARMORSTEIN, R. & HARRISON, S. C. 1994. Crystal structure of a PPR1-DNA complex: DNA recognition by proteins containing a Zn2Cys6 binuclear cluster. *Genes Dev,* 8**,** 2504-12.

MASKELL, D. P., HU, X. W. & SINGLETON, M. R. 2010. Molecular architecture and assembly of the yeast kinetochore MIND complex. *J Cell Biol,* 190**,** 823-34.

MCAINSH, A. D., TYTELL, J. D. & SORGER, P. K. 2003. Structure, function, and regulation of budding yeast kinetochores. *Annu Rev Cell Dev Biol,* 19**,** 519-39.

MCINTOSH, J. R. 1991. Structural and mechanical control of mitotic progression. *Cold Spring Harb Symp Quant Biol,* 56**,** 613-9.

MCMULLAN, G., FARUQI, A. R., CLARE, D. & HENDERSON, R. 2014. Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy,* 147**,** 156-63.

MEASDAY, V., HAILEY, D. W., POT, I., GIVAN, S. A., HYLAND, K. M., CAGNEY, G., FIELDS, S., DAVIS, T. N. & HIETER, P. 2002. Ctf3p, the Mis6 budding yeast homolog, interacts with Mcm22p and Mcm16p at the yeast outer kinetochore. *Genes Dev,* 16**,** 101-13.

MELLOR, J., JIANG, W., FUNK, M., RATHJEN, J., BARNES, C. A., HINZ, T., HEGEMANN, J. H. & PHILIPPSEN, P. 1990. CPF1, a yeast protein which functions in centromeres and promoters. *EMBO J,* 9**,** 4017-26.

MELUH, P. B. & KOSHLAND, D. 1995. Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. *Mol Biol Cell,* 6**,** 793-807.

MELUH, P. B. & KOSHLAND, D. 1997. Budding yeast centromere composition and assembly as revealed by in vivo cross-linking. *Genes Dev,* 11**,** 3401-12.

MELUH, P. B., YANG, P., GLOWCZEWSKI, L., KOSHLAND, D. & SMITH, M. M. 1998. Cse4p is a component of the core centromere of Saccharomyces cerevisiae. *Cell,* 94**,** 607-13.

MERALDI, P., MCAINSH, A. D., RHEINBAY, E. & SORGER, P. K. 2006. Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. *Genome Biol,* 7**,** R23.

MINDELL, J. A. & GRIGORIEFF, N. 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol,* 142**,** 334-47.

MINE-HATTAB, J. & ROTHSTEIN, R. 2012. Increased chromosome mobility facilitates homology search during recombination. *Nat Cell Biol,* 14**,** 510-7.

MIZUGUCHI, G., XIAO, H., WISNIEWSKI, J., SMITH, M. M. & WU, C. 2007. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell,* 129**,** 1153-64.

MORENO-MORENO, O., TORRAS-LLORT, M. & AZORIN, F. 2017. Variations on a nucleosome theme: The structural basis of centromere function. *Bioessays*.

MURATA-HORI, M. & WANG, Y. L. 2002. The kinase activity of aurora B is required for kinetochore-microtubule interactions during mitosis. *Curr Biol,* 12**,** 894-9.

MUSACCHIO, A. 2015. The Molecular Biology of Spindle Assembly Checkpoint Signaling Dynamics. *Curr Biol,* 25**,** R1002-18.

MUSACCHIO, A. & DESAI, A. 2017. A Molecular View of Kinetochore Assembly and Function. *Biology (Basel),* 6.

NARDI, I. K., ZASADZINSKA, E., STELLFOX, M. E., KNIPPLER, C. M. & FOLTZ, D. R. 2016. Licensing of Centromeric Chromatin Assembly through the Mis18alpha-Mis18beta Heterotetramer. *Mol Cell,* 61**,** 774-87.

NG, R. & CARBON, J. 1987. Mutational and in vitro protein-binding studies on centromere DNA from Saccharomyces cerevisiae. *Mol Cell Biol,* 7**,** 4522-34.

NICKLAS, R. B. 1997. How cells get the right chromosomes. *Science,* 275**,** 632-7.

NICKLAS, R. B. & KOCH, C. A. 1969. Chromosome micromanipulation. 3. Spindle fiber tension and the reorientation of mal-oriented chromosomes. *J Cell Biol,* 43**,** 40-50.

NICKLAS, R. B. & WARD, S. C. 1994. Elements of error correction in mitosis: microtubule capture, release, and tension. *J Cell Biol,* 126**,** 1241-53.

NIEDENTHAL, R. K., SEN-GUPTA, M., WILMEN, A. & HEGEMANN, J. H. 1993. Cpf1 protein induced bending of yeast centromere DNA element I. *Nucleic Acids Res,* 21**,** 4726-33.

NILSSON, J., YEKEZARE, M., MINSHULL, J. & PINES, J. 2008. The APC/C maintains the spindle assembly checkpoint by targeting Cdc20 for destruction. *Nat Cell Biol,* 10**,** 1411-20.

NISHINO, T., RAGO, F., HORI, T., TOMII, K., CHEESEMAN, I. M. & FUKAGAWA, T. 2013. CENP-T provides a structural platform for outer kinetochore assembly. *EMBO J,* 32**,** 424-36.

NISHINO, T., TAKEUCHI, K., GASCOIGNE, K. E., SUZUKI, A., HORI, T., OYAMA, T., MORIKAWA, K., CHEESEMAN, I. M. & FUKAGAWA, T. 2012. CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. *Cell,* 148**,** 487-501.

NOGALES, E. 2018. Profile of Joachim Frank, Richard Henderson, and Jacques Dubochet, 2017 Nobel Laureates in Chemistry. *Proc Natl Acad Sci U S A,* 115**,** 441-444.

OHI, M., LI, Y., CHENG, Y. & WALZ, T. 2004. Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. *Biol Proced Online,* 6**,** 23-34.

OIKONOMOU, C. M., CHANG, Y. W. & JENSEN, G. J. 2016. A new view into prokaryotic cell biology from electron cryotomography. *Nat Rev Microbiol,* 14**,** 205-20.

OKADA, T., OHZEKI, J., NAKANO, M., YODA, K., BRINKLEY, W. R., LARIONOV, V. & MASUMOTO, H. 2007. CENP-B controls centromere formation depending on the chromatin context. *Cell,* 131**,** 1287-300.

ORLICKY, S., TANG, X., NEDUVA, V., ELOWE, N., BROWN, E. D., SICHERI, F. & TYERS, M. 2010. An allosteric inhibitor of substrate recognition by the SCF(Cdc4) ubiquitin ligase. *Nat Biotechnol,* 28**,** 733-7.

ORLICKY, S., TANG, X., WILLEMS, A., TYERS, M. & SICHERI, F. 2003. Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. *Cell,* 112**,** 243-56.

ORLOVA, E. V. & SAIBIL, H. R. 2011. Structural analysis of macromolecular assemblies by electron microscopy. *Chem Rev,* 111**,** 7710-48.

ORTIZ, J., STEMMANN, O., RANK, S. & LECHNER, J. 1999. A putative protein complex consisting of Ctf19, Mcm21, and Okp1 represents a missing link in the budding yeast kinetochore. *Genes Dev,* 13**,** 1140-55.

ÖSTERGREN, G. 1951. *The mechanism of co-orientation in bivalents, and multivalents,* Lund,, Berlingska boktr.

PALMER, D. K., O'DAY, K., WENER, M. H., ANDREWS, B. S. & MARGOLIS, R. L. 1987. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J Cell Biol,* 104**,** 805-15.

PASSMORE, L. A. & RUSSO, C. J. 2016. Specimen Preparation for High-Resolution Cryo-EM. *Methods Enzymol,* 579**,** 51-86.

PEKGOZ ALTUNKAYA, G., MALVEZZI, F., DEMIANOVA, Z., ZIMNIAK, T., LITOS, G., WEISSMANN, F., MECHTLER, K., HERZOG, F. & WESTERMANN, S. 2016. CCAN Assembly Configures Composite Binding Interfaces to Promote Cross-Linking of Ndc80 Complexes at the Kinetochore. *Curr Biol,* 26**,** 2370-8.

PENCZEK, P. A., ZHU, J., SCHRÖDER, R. & FRANK, J. 1997. THREE DIMENSIONAL RECONSTRUCTION WITH CONTRAST TRANSFER COMPENSATION FROM DEFOCUS SERIES. *Scanning Microscopy* 11**,** 147-154.

PERPELESCU, M. & FUKAGAWA, T. 2011. The ABCs of CENPs. *Chromosoma,* 120**,** 425-46.

PERRICHES, T. 2014. *The CBF3 complex structure and function during point centromere establishment.* PhD, University College London.

PERRICHES, T. & SINGLETON, M. R. 2012. Structure of yeast kinetochore Ndc10 DNA-binding domain reveals unexpected evolutionary relationship to tyrosine recombinases. *J Biol Chem,* 287**,** 5173-9.

PETROVIC, A., KELLER, J., LIU, Y., OVERLACK, K., JOHN, J., DIMITROVA, Y. N., JENNI, S., VAN GERWEN, S., STEGE, P., WOHLGEMUTH, S., ROMBAUT, P., HERZOG, F., HARRISON, S. C., VETTER, I. R. & MUSACCHIO, A. 2016. Structure of the MIS12 Complex and Molecular Basis of Its Interaction with CENP-C at Human Kinetochores. *Cell,* 167**,** 1028-1040 e15.

PETROVIC, A., MOSALAGANTI, S., KELLER, J., MATTIUZZO, M., OVERLACK, K., KRENN, V., DE ANTONI, A., WOHLGEMUTH, S., CECATIELLO, V., PASQUALATO, S., RAUNSER, S. & MUSACCHIO, A. 2014. Modular assembly of RWD domains on the Mis12 complex underlies outer kinetochore organization. *Mol Cell,* 53**,** 591-605.

PETROVIC, A., PASQUALATO, S., DUBE, P., KRENN, V., SANTAGUIDA, S., CITTARO, D., MONZANI, S., MASSIMILIANO, L., KELLER, J., TARRICONE,

A., MAIOLICA, A., STARK, H. & MUSACCHIO, A. 2010. The MIS12 complex is a protein interaction hub for outer kinetochore assembly. *J Cell Biol,* 190**,** 835-52.

PETTERSEN, E. F., GODDARD, T. D., HUANG, C. C., COUCH, G. S., GREENBLATT, D. M., MENG, E. C. & FERRIN, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem,* 25**,** 1605-12.

PIETRASANTA, L. I. T., D.; HSIEH, W.; RAO, S.; STEMMANN, O.; LECHNER, J.; CARBON, J., HANSMA, H. 1999. Probing the Saccharomyces cerevisiae centromeric DNA (CEN DNA)–binding factor 3 (CBF3) kinetochore complex by using atomic force microscopy. *Proc Natl Acad Sci U S A,* 96**,** 3757-3762.

PRIMORAC, I., WEIR, J. R., CHIROLI, E., GROSS, F., HOFFMANN, I., VAN GERWEN, S., CILIBERTO, A. & MUSACCHIO, A. 2013. Bub3 reads phosphorylated MELT repeats to promote spindle assembly checkpoint signaling. *Elife,* 2**,** e01030.

PUNJANI, A., RUBINSTEIN, J. L., FLEET, D. J. & BRUBAKER, M. A. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods,* 14**,** 290-296.

PURVIS, A. & SINGLETON, M. R. 2008. Insights into kinetochore-DNA interactions from the structure of Cep3Delta. *EMBO Rep,* 9**,** 56-62.

REITSMA, J. M., LIU, X., REICHERMEIER, K. M., MORADIAN, A., SWEREDOSKI, M. J., HESS, S. & DESHAIES, R. J. 2017. Composition and Regulation of the Cellular Repertoire of SCF Ubiquitin Ligases. *Cell,* 171**,** 1326-1339 e14.

ROBERT, X. & GOUET, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res,* 42**,** W320-4.

RODRIGO-BRENNI, M. C. T., S.; BOUCK, D.C.; KAPLAN, K.B. 2004. Sgt1p and Skp1p Modulate the Assembly and Turnover of CBF3 Complexes Required for Proper Kinetochore Function. *Mol Biol Cell,* 15**,** 3366-3378.

ROHOU, A. & GRIGORIEFF, N. 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol,* 192**,** 216-21.

ROSENBERG, J. S., CROSS, F. R. & FUNABIKI, H. 2011. KNL1/Spc105 recruits PP1 to silence the spindle assembly checkpoint. *Curr Biol,* 21**,** 942-7.

ROSENTHAL, P. B. & HENDERSON, R. 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol,* 333**,** 721-45.

RUSKIN, R. S., YU, Z. & GRIGORIEFF, N. 2013. Quantitative characterization of electron detectors for transmission electron microscopy. *J Struct Biol,* 184**,** 385-93.

RUSSELL, I. D. G., A.S.; SORGER, P.K. 1999. The Unstable F-box Protein p58-Ctf13 Forms the Structural Core of the

CBF3 Kinetochore Complex. *J Cell Biol,* 145**,** 933-950.

SANCHEZ-PULIDO, L., PIDOUX, A. L., PONTING, C. P. & ALLSHIRE, R. C. 2009. Common ancestry of the CENP-A chaperones Scm3 and HJURP. *Cell,* 137**,** 1173-4.

SANTAGUIDA, S. & MUSACCHIO, A. 2009. The life and miracles of kinetochores. *EMBO J,* 28**,** 2511-31.

SARANGAPANI, K. K., AKIYOSHI, B., DUGGAN, N. M., BIGGINS, S. & ASBURY, C. L. 2013. Phosphoregulation promotes release of kinetochores from dynamic microtubules via multiple mechanisms. *Proc Natl Acad Sci U S A,* 110**,** 7282-7.

SCHERES, S. H. 2012. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol,* 180**,** 519-30.

SCHERES, S. H. 2014. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife,* 3**,** e03665.

SCHERES, S. H. 2016. Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods Enzymol,* 579**,** 125-57.
SCHITTENHELM, R. B., CHALECKIS, R. & LEHNER, C. F. 2009. Intrakinetochore localization and essential functional domains of Drosophila Spc105. *EMBO J,* 28**,** 2374-86.
SCHLEIFFER, A., MAIER, M., LITOS, G., LAMPERT, F., HORNUNG, P., MECHTLER, K. & WESTERMANN, S. 2012. CENP-T proteins are conserved centromere receptors of the Ndc80 complex. *Nat Cell Biol,* 14**,** 604-13.
SCHMIDT, J. C., ARTHANARI, H., BOESZOERMENYI, A., DASHKEVICH, N. M., WILSON-KUBALEK, E. M., MONNIER, N., MARKUS, M., OBERER, M., MILLIGAN, R. A., BATHE, M., WAGNER, G., GRISHCHUK, E. L. & CHEESEMAN, I. M. 2012. The kinetochore-bound Ska1 complex tracks depolymerizing microtubules and binds to curved protofilaments. *Dev Cell,* 23**,** 968-80.
SCHMITZBERGER, F. & HARRISON, S. C. 2012. RWD domain: a recurring module in kinetochore architecture shown by a Ctf19-Mcm21 complex structure. *EMBO Rep,* 13**,** 216-22.
SCHMITZBERGER, F., RICHTER, M. M., GORDIYENKO, Y., ROBINSON, C. V., DADLEZ, M. & WESTERMANN, S. 2017. Molecular basis for inner kinetochore configuration through RWD domain-peptide interactions. *EMBO J,* 36**,** 3458-3482.
SCHULMAN, B. A., CARRANO, A. C., JEFFREY, P. D., BOWEN, Z., KINNUCAN, E. R., FINNIN, M. S., ELLEDGE, S. J., HARPER, J. W., PAGANO, M. & PAVLETICH, N. P. 2000. Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature,* 408**,** 381-6.
SCHUYLER, S. C., WU, Y. F. & KUAN, V. J. 2012. The Mad1-Mad2 balancing act--a damaged spindle checkpoint in chromosome instability and cancer. *J Cell Sci,* 125**,** 4197-206.
SCHWAB, M., LUTUM, A. S. & SEUFERT, W. 1997. Yeast Hct1 is a regulator of Clb2 cyclin proteolysis. *Cell,* 90**,** 683-93.
SEOL, J. H., SHEVCHENKO, A., SHEVCHENKO, A. & DESHAIES, R. J. 2001. Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. *Nat Cell Biol,* 3**,** 384-91.
SHIMOGAWA, M. M., WIDLUND, P. O., RIFFLE, M., ESS, M. & DAVIS, T. N. 2009. Bir1 is required for the tension checkpoint. *Mol Biol Cell,* 20**,** 915-23.
SHUAIB, M., OUARARHNI, K., DIMITROV, S. & HAMICHE, A. 2010. HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. *Proc Natl Acad Sci U S A,* 107**,** 1349-54.
SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SODING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol,* 7**,** 539.
SINGH, T. R., SARO, D., ALI, A. M., ZHENG, X. F., DU, C. H., KILLEN, M. W., SACHPATZIDIS, A., WAHENGBAM, K., PIERCE, A. J., XIONG, Y., SUNG, P. & MEETEI, A. R. 2010. MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Mol Cell,* 37**,** 879-86.
SIVAKUMAR, S., JANCZYK, P. L., QU, Q., BRAUTIGAM, C. A., STUKENBERG, P. T., YU, H. & GORBSKY, G. J. 2016. The human SKA complex drives the metaphase-anaphase cell cycle transition by recruiting protein phosphatase 1 to kinetochores. *Elife,* 5.
SMITH, D. J. 2008. Ultimate resolution in the electron microscope? *MaterialsToday,* 11**,** 30-38.

SORGER, P. K., DOHENY, K. F., HIETER, P., KOPSKI, K. M., HUFFAKER, T. C. & HYMAN, A. A. 1995. Two genes required for the binding of an essential Saccharomyces cerevisiae kinetochore complex to DNA. *Proc Natl Acad Sci U S A,* 92**,** 12026-30.

SORGER, P. K. S., F.E.; HYMAN, A.A. 1994. Factors Required for the Binding of Reassembled Yeast Kinetochores to Microtubules In Vitro. *J Cell Biol,* 127**,** 995-1008.

SPENCER, F. & HIETER, P. 1992. Centromere DNA mutations induce a mitotic delay in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A,* 89**,** 8908-12.

STEINER, F. A. & HENIKOFF, S. 2015. Diversity in the organization of centromeric chromatin. *Curr Opin Genet Dev,* 31**,** 28-35.

STELLFOX, M. E., BAILEY, A. O. & FOLTZ, D. R. 2013. Putting CENP-A in its place. *Cell Mol Life Sci,* 70**,** 387-406.

STEMMANN, O. & LECHNER, J. 1996. The Saccharomyces cerevisiae kinetochore contains a cyclin-CDK complexing homologue, as identified by in vitro reconstitution. *EMBO J,* 15**,** 3611-20.

STEMMANN, O., NEIDIG, A., KOCHER, T., WILM, M. & LECHNER, J. 2002. Hsp90 enables Ctf13p/Skp1p to nucleate the budding yeast kinetochore. *Proc Natl Acad Sci U S A,* 99**,** 8585-90.

STOLER, S., KEITH, K. C., CURNICK, K. E. & FITZGERALD-HAYES, M. 1995. A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes Dev,* 9**,** 573-86.

STOLER, S., ROGERS, K., WEITZE, S., MOREY, L., FITZGERALD-HAYES, M. & BAKER, R. E. 2007. Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. *Proc Natl Acad Sci U S A,* 104**,** 10571-6.

STRECKER, J., GUPTA, G. D., ZHANG, W., BASHKUROV, M., LANDRY, M. C., PELLETIER, L. & DUROCHER, D. 2016. DNA damage signalling targets the kinetochore to promote chromatin mobility. *Nat Cell Biol,* 18**,** 281-90.

STRUNNIKOV, A. V., KINGSBURY, J. & KOSHLAND, D. 1995. CEP3 encodes a centromere protein of Saccharomyces cerevisiae. *J Cell Biol,* 128**,** 749-60.

STUTTMANN, J., PARKER, J. E. & NOEL, L. D. 2008. Staying in the fold: The SGT1/chaperone machinery in maintenance and evolution of leucine-rich repeat proteins. *Plant Signal Behav,* 3**,** 283-5.

SUBRAMANIAM, S., KUHLBRANDT, W. & HENDERSON, R. 2016. CryoEM at IUCrJ: a new era. *IUCrJ,* 3**,** 3-7.

SUDAKIN, V., CHAN, G. K. & YEN, T. J. 2001. Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. *J Cell Biol,* 154**,** 925-36.

SUDAKIN, V., GANOTH, D., DAHAN, A., HELLER, H., HERSHKO, J., LUCA, F. C., RUDERMAN, J. V. & HERSHKO, A. 1995. The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol Biol Cell,* 6**,** 185-97.

SULLIVAN, K. F., HECHENBERGER, M. & MASRI, K. 1994. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J Cell Biol,* 127**,** 581-92.

SUMA, M., KITAGAWA, T., NAKASE, Y., NAKAZAWA, N., YANAGIDA, M. & MATSUMOTO, T. 2018. Fission Yeast CENP-C (Cnp3) Plays a Role in Restricting the Site of CENP-A Accumulation. *G3 (Bethesda)*.

SUZUKI, A., BADGER, B. L., HAASE, J., OHASHI, T., ERICKSON, H. P., SALMON, E. D. & BLOOM, K. 2016. How the kinetochore couples microtubule force and centromere stretch to move chromosomes. *Nat Cell Biol,* 18**,** 382-92.

TACHIWANA, H., KAGAWA, W., SHIGA, T., OSAKABE, A., MIYA, Y., SAITO, K., HAYASHI-TAKANAKA, Y., ODA, T., SATO, M., PARK, S. Y., KIMURA, H. & KURUMIZAKA, H. 2011. Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature,* 476**,** 232-5.

TAIPALE, M., TUCKER, G., PENG, J., KRYKBAEVA, I., LIN, Z. Y., LARSEN, B., CHOI, H., BERGER, B., GINGRAS, A. C. & LINDQUIST, S. 2014. A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell,* 158**,** 434-448.

TAN, Y. Z., BALDWIN, P. R., DAVIS, J. H., WILLIAMSON, J. R., POTTER, C. S., CARRAGHER, B. & LYUMKIS, D. 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods,* 14**,** 793-796.

TAN, Y. Z., CHENG, A., POTTER, C. S. & CARRAGHER, B. 2016. Automated data collection in single particle electron microscopy. *Microscopy (Oxf),* 65**,** 43-56.

TANAKA, K. & HIROTA, T. 2009. Chromosome segregation machinery and cancer. *Cancer Sci,* 100**,** 1158-65.

TANAKA, Y., NUREKI, O., KURUMIZAKA, H., FUKAI, S., KAWAGUCHI, S., IKUTA, M., IWAHARA, J., OKAZAKI, T. & YOKOYAMA, S. 2001. Crystal structure of the CENP-B protein-DNA complex: the DNA-binding domains of CENP-B induce kinks in the CENP-B box DNA. *EMBO J,* 20**,** 6612-8.

TANG, G., PENG, L., BALDWIN, P. R., MANN, D. S., JIANG, W., REES, I. & LUDTKE, S. J. 2007. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol,* 157**,** 38-46.

TANG, X., ORLICKY, S., MITTAG, T., CSIZMOK, V., PAWSON, T., FORMAN-KAY, J. D., SICHERI, F. & TYERS, M. 2012. Composite low affinity interactions dictate recognition of the cyclin-dependent kinase inhibitor Sic1 by the SCFCdc4 ubiquitin ligase. *Proc Natl Acad Sci U S A,* 109**,** 3287-92.

TAVORMINA, P. A. & BURKE, D. J. 1998. Cell cycle arrest in cdc20 mutants of Saccharomyces cerevisiae is independent of Ndc10p and kinetochore function but requires a subset of spindle checkpoint genes. *Genetics,* 148**,** 1701-13.

THOMAS, S. & KAPLAN, K. B. 2007. A Bir1p Sli15p kinetochore passenger complex regulates septin organization during anaphase. *Mol Biol Cell,* 18**,** 3820-34.

TIEN, J. F., UMBREIT, N. T., GESTAUT, D. R., FRANCK, A. D., COOPER, J., WORDEMAN, L., GONEN, T., ASBURY, C. L. & DAVIS, T. N. 2010. Cooperation of the Dam1 and Ndc80 kinetochore complexes enhances microtubule coupling and is regulated by aurora B. *J Cell Biol,* 189**,** 713-23.

UMBREIT, N. T., GESTAUT, D. R., TIEN, J. F., VOLLMAR, B. S., GONEN, T., ASBURY, C. L. & DAVIS, T. N. 2012. The Ndc80 kinetochore complex directly modulates microtubule dynamics. *Proc Natl Acad Sci U S A,* 109**,** 16113-8.

VAN DEN ENT, F. & LOWE, J. 2006. RF cloning: a restriction-free method for inserting target genes into plasmids. *J Biochem Biophys Methods,* 67**,** 67-74.

VLIJM, R., KIM, S. H., DE ZWART, P. L., DALAL, Y. & DEKKER, C. 2017. The supercoiling state of DNA determines the handedness of both H3 and CENP-A nucleosomes. *Nanoscale*.

VONCK, J. & MILLS, D. J. 2017. Advances in high-resolution cryo-EM of oligomeric enzymes. *Curr Opin Struct Biol,* 46**,** 48-54.

WAN, X., O'QUINN, R. P., PIERCE, H. L., JOGLEKAR, A. P., GALL, W. E., DELUCA, J. G., CARROLL, C. W., LIU, S. T., YEN, T. J., MCEWEN, B. F., STUKENBERG, P. T., DESAI, A. & SALMON, E. D. 2009. Protein architecture of the human kinetochore microtubule attachment site. *Cell,* 137**,** 672-84.

WANG, F., DAI, J., DAUM, J. R., NIEDZIALKOWSKA, E., BANERJEE, B., STUKENBERG, P. T., GORBSKY, G. J. & HIGGINS, J. M. 2010. Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. *Science,* 330**,** 231-5.

WANG, H. W., LONG, S., CIFERRI, C., WESTERMANN, S., DRUBIN, D., BARNES, G. & NOGALES, E. 2008. Architecture and flexibility of the yeast Ndc80 kinetochore complex. *J Mol Biol,* 383**,** 894-903.

WEI, R. R., AL-BASSAM, J. & HARRISON, S. C. 2007. The Ndc80/HEC1 complex is a contact point for kinetochore-microtubule attachment. *Nat Struct Mol Biol,* 14**,** 54-9.

WEI, R. R., SORGER, P. K. & HARRISON, S. C. 2005. Molecular organization of the Ndc80 complex, an essential kinetochore component. *Proc Natl Acad Sci U S A,* 102**,** 5363-7.

WEIR, J. R., FAESEN, A. C., KLARE, K., PETROVIC, A., BASILICO, F., FISCHBOCK, J., PENTAKOTA, S., KELLER, J., PESENTI, M. E., PAN, D., VOGT, D., WOHLGEMUTH, S., HERZOG, F. & MUSACCHIO, A. 2016. Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature,* 537**,** 249-253.

WELBURN, J. P., VLEUGEL, M., LIU, D., YATES, J. R., 3RD, LAMPSON, M. A., FUKAGAWA, T. & CHEESEMAN, I. M. 2010. Aurora B phosphorylates spatially distinct targets to differentially regulate the kinetochore-microtubule interface. *Mol Cell,* 38**,** 383-92.

WESTHORPE, F. G. & STRAIGHT, A. F. 2013. Functions of the centromere and kinetochore in chromosome segregation. *Curr Opin Cell Biol,* 25**,** 334-40.

WIELAND, G., ORTHAUS, S., OHNDORF, S., DIEKMANN, S. & HEMMERICH, P. 2004. Functional complementation of human centromere protein A (CENP-A) by Cse4p from Saccharomyces cerevisiae. *Mol Cell Biol,* 24**,** 6620-30.

WILLHOFT, O., KERR, R., PATEL, D., ZHANG, W., AL-JASSAR, C., DAVITER, T., MILLSON, S. H., THALASSINOS, K. & VAUGHAN, C. K. 2017. The crystal structure of the Sgt1-Skp1 complex: the link between Hsp90 and both SCF E3 ubiquitin ligases and kinetochores. *Sci Rep,* 7**,** 41626.

WONG, K., PERPICH, J. D., KOZLOV, G., CYGLER, M., ABU KWAIK, Y. & GEHRING, K. 2017. Structural Mimicry by a Bacterial F Box Effector Hijacks the Host Ubiquitin-Proteasome System. *Structure,* 25**,** 376-383.

WU, S., ARMACHE, J. P. & CHENG, Y. 2016. Single-particle cryo-EM data acquisition by using direct electron detection camera. *Microscopy (Oxf),* 65**,** 35-41.

XIAO, H., WANG, F., WISNIEWSKI, J., SHAYTAN, A. K., GHIRLANDO, R., FITZGERALD, P. C., HUANG, Y., WEI, D., LI, S., LANDSMAN, D., PANCHENKO, A. R. & WU, C. 2017. Molecular basis of CENP-C association with the CENP-A nucleosome at yeast centromeres. *Genes Dev*.

YAMAMOTO, A., GUACCI, V. & KOSHLAND, D. 1996. Pds1p, an inhibitor of anaphase in budding yeast, plays a critical role in the APC and checkpoint pathway(s). *J Cell Biol,* 133**,** 99-110.

YANG, Y., WU, F., WARD, T., YAN, F., WU, Q., WANG, Z., MCGLOTHEN, T., PENG, W., YOU, T., SUN, M., CUI, T., HU, R., DOU, Z., ZHU, J., XIE, W., RAO, Z., DING, X. & YAO, X. 2008. Phosphorylation of HsMis13 by Aurora B kinase is essential for assembly of functional kinetochore. *J Biol Chem,* 283**,** 26726-36.

YODA, K., ANDO, S., MORISHITA, S., HOUMURA, K., HASHIMOTO, K., TAKEYASU, K. & OKAZAKI, T. 2000. Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. *Proc Natl Acad Sci U S A,* 97**,** 7266-71.

YOON, H. J. & CARBON, J. 1999. Participation of Bir1p, a member of the inhibitor of apoptosis family, in yeast chromosome segregation events. *Proc Natl Acad Sci U S A,* 96**,** 13208-13.

ZAYTSEV, A. V., MICK, J. E., MASLENNIKOV, E., NIKASHIN, B., DELUCA, J. G. & GRISHCHUK, E. L. 2015. Multisite phosphorylation of the NDC80 complex gradually tunes its microtubule-binding affinity. *Mol Biol Cell,* 26**,** 1829-44.

ZAYTSEV, A. V., SUNDIN, L. J., DELUCA, K. F., GRISHCHUK, E. L. & DELUCA, J. G. 2014. Accurate phosphoregulation of kinetochore-microtubule affinity requires unconstrained molecular interactions. *J Cell Biol,* 206**,** 45-59.

ZHANG, K. 2016. Gctf: Real-time CTF determination and correction. *J Struct Biol,* 193**,** 1-12.

ZHANG, W., LUKOYNOVA, N., MIAH, S., LUCAS, J. & VAUGHAN, C. K. 2018. Insights into Centromere DNA Bending Revealed by the Cryo-EM Structure of the Core Centromere Binding Factor 3 with Ndc10. *Cell Rep,* 24**,** 744-754.

ZHENG, L., CHEN, Y. & LEE, W. H. 1999. Hec1p, an evolutionarily conserved coiled-coil protein, modulates chromosome segregation through interaction with SMC proteins. *Mol Cell Biol,* 19**,** 5417-28.

ZHENG, N., SCHULMAN, B. A., SONG, L., MILLER, J. J., JEFFREY, P. D., WANG, P., CHU, C., KOEPP, D. M., ELLEDGE, S. J., PAGANO, M., CONAWAY, R. C., CONAWAY, J. W., HARPER, J. W. & PAVLETICH, N. P. 2002. Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature,* 416**,** 703-9.

ZHENG, S. Q., PALOVCAK, E., ARMACHE, J. P., VERBA, K. A., CHENG, Y. & AGARD, D. A. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods,* 14**,** 331-332.

ZHOU, Z., FENG, H., ZHOU, B. R., GHIRLANDO, R., HU, K., ZWOLAK, A., MILLER JENKINS, L. M., XIAO, H., TJANDRA, N., WU, C. & BAI, Y. 2011. Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. *Nature,* 472**,** 234-7.

ZHU, D., WANG, X., FANG, Q., VAN ETTEN, J. L., ROSSMANN, M. G., RAO, Z. & ZHANG, X. 2018. Pushing the resolution limit by correcting the Ewald sphere effect in single-particle Cryo-EM reconstructions. *Nat Commun,* 9**,** 1552.

ZINKOWSKI, R. P., MEYNE, J. & BRINKLEY, B. R. 1991. The centromere-kinetochore complex: a repeat subunit model. *J Cell Biol,* 113**,** 1091-110.